




Characterizing multi-decadal, annual land cover change dynamics in Houston, TX based on automated classification of Landsat imagery

C.R. Hakkenberg ^a, M.P. Dannenberg^{b,c}, C. Song^d and K.B. Ensor^a

^aDepartment of Statistics, Rice University, Houston, TX, USA; ^bDepartment of Geographical and Sustainability Sciences, University of Iowa, Iowa City, IA, USA; ^cSchool of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA; ^dDepartment of Geography, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

ABSTRACT

In 2017, Hurricane Harvey caused substantial loss of life and property in the swiftly urbanizing region of Houston, TX. Now in its wake, researchers are tasked with investigating how to plan for and mitigate the impact of similar events in the future, despite expectations of increased storm intensity and frequency as well as accelerating urbanization trends. Critical to this task is the development of automated workflows for producing accurate and consistent land cover maps of sufficiently fine spatio-temporal resolution over large areas and long timespans. In this study, we developed an innovative automated classification algorithm that overcomes some of the traditional trade-offs between fine spatio-temporal resolution and extent – to produce a multi-scene, 30m annual land cover time series characterizing 21 years of land cover dynamics in the 35,000 km² Greater Houston area. The ensemble algorithm takes advantage of the synergistic value of employing all acceptable Landsat imagery in a given year, using aggregate votes from the posterior predictive distributions of multiple image composites to mitigate against misclassifications in any one image, and fill gaps due to missing and contaminated data, such as those from clouds and cloud shadows. The procedure is fully automated, combining adaptive signature generalization and spatio-temporal stabilization for consistency across sensors and scenes. The land cover time series is validated using independent, multi-temporal fine-resolution imagery, achieving crisp overall accuracies between 78–86% and fuzzy overall accuracies between 91–94%. Validated maps and corresponding areal cover estimates corroborate what census and economic data from the Greater Houston area likewise indicate: rapid growth from 1997–2017, demonstrated by the conversion of 2,040 km² (± 400 km²) to developed land cover, 14% of which resulted from the conversion of wetlands. Beyond its implications for urbanization trends in Greater Houston, this study demonstrates the potential for automated approaches to quantifying large extent, fine resolution land cover change, as well as the added value of temporally-dense time series for characterizing higher-order spatio-temporal dynamics of land cover, including periodicity, abrupt transitions, and time lags from underlying demographic and socio-economic trends.

ARTICLE HISTORY

Received 15 June 2018
Accepted 18 August 2018

CONTACT C.R. Hakkenberg  ch55@rice.edu  Department of Statistics, Rice University, Duncan Hall #2077, Houston, TX 77251, USA

 Supplementary data for this article can be accessed [here](#).

© 2018 Informa UK Limited, trading as Taylor & Francis Group

1. Introduction

When Hurricane Harvey made landfall in Texas in August 2017, it resulted in the largest rainfall event on record in the US, producing as much as 1200 mm of rain over a seven-day period. The hurricane and subsequent flooding resulted in at least 89 deaths, 30,000 displaced people, and \$125 billion dollars in damage – its impact exacerbated as stalled over one of the US's largest urban areas: Houston, TX (NOAA 2018). Over the past several decades of rapid growth and development, Greater Houston has adopted a resistance-based flood risk reduction strategy, relying on large-scale engineering solutions to distribute the increased run-off associated with its large-scale, largely-unzoned urban development (Brody, Kim, and Gunn 2013). However, despite these infrastructural improvements, flood vulnerability persists due in part to the vast expansion of low intensity impervious land cover characteristic of sprawling urbanization (Jaret et al. 2009). Owing to the simultaneous expectation of higher frequency and stronger intensity hurricanes in the region (Knutson et al. 2010; Emanuel 2017), studies are urgently needed to investigate the independent and interactive aspects of global climate change and local land cover conversion in contributing to storm damage across vulnerable urban areas like Houston. Critical to this effort is the development of automated workflows for producing accurate and consistent land cover maps capable of characterizing historical patterns and temporal trajectories of land cover change, as well as their spatially-variant change rates at a sufficiently fine spatio-temporal resolution.

In this regard, the Landsat satellite data archive offers researchers an unparalleled source of historical medium resolution optical imagery, enabling the compilation of multi-decadal land-cover change trajectories – temporal sequences of land-cover classes derived from satellite images at multiple dates (Loveland and Dwyer 2012; Gómez, White, and Wulder 2016). However, the production of annual land cover classifications over multiple Landsat scene extents and long timespans is complicated by a number of factors including radiometric inconsistencies in reflectance retrievals through space (between neighbouring paths) and time (between sensors) (Vogelmann et al. 2016). In addition, low acquisition frequency may result in irregular dates of usable imagery, exacerbating differences in scene conditions due to changes in land surface phenology, atmospheric conditions, and illumination angles (C. Song et al. 2015; Song and Woodcock 2003). These considerations have led some researchers to employ multi-year imagery for classification surrounding a nominal year, resulting in a sparse time series at a frequency on the order of 7–10 years (Sexton et al. 2013; Fenta et al. 2017) to 3–6 years (Dou and Chen 2017; Homer et al. 2015). And while the expanded temporal window for input imagery often results in high quality map products, they may not be precise enough to accurately reflect land cover conditions for the nominal year and, as a time series, may be too coarse to capture higher-order temporal dynamics critical to assessing spatio-temporal complexities of human–environment systems (Jensen and Cowen 1999; Lunetta et al. 2004). In response, recent studies have focused on a range of data fusion, composite, and interpolation approaches to create land cover time series at increasingly fine resolutions and large extents in the spatial and temporal domains (Gong et al. 2013; Song et al. 2016; Li, Gong, and Liang 2015).

In this study, we present a multi-scene, annual land cover time series characterizing 21 years of land cover trends in the 35,000 km² Greater Houston area. The methodology

employed is unique in that it entirely automates the image processing, data fusion, and classification workflow to produce a temporally dense and consistent land cover time series using adaptive signature generalization, multi-scene compositing and ensemble classification using all acceptable Landsat imagery in a given year, as well as spatio-temporal stabilization for consistency across sensors and scenes. A distinct merit of this study is that the proposed automated classification method overcomes some of the traditional trade-offs between spatio-temporal resolution and extent, with final maps possessing a fine resolution (annual, 30m) over a large duration and extent (21 years, 35,000 km²). The resulting map time series is compared with concurrent NLCD products, and validated using multi-year, fine resolution independent reference imagery. Using results from the probability-based sampling design of the accuracy assessment procedure, we quantify the areal extent of land cover conversions, as well as change rates. As a case study quantifying the rapid urbanization of Greater Houston, this research demonstrates the potential for automated remote sensing workflows to move beyond bi-temporal change detection to characterize higher-order, annual spatio-temporal dynamics of land cover change, including periodicity, abrupt transitions, and time lags emerging from underlying demographic and socio-economic trends.

2. Materials and methods

2.1. Study area

The 35,000 km² study area consists of the 13 counties defining the Houston-Galveston Area (HGAC 2018), namely: Austin, Brazoria, Chambers, Colorado, Fort Bend, Galveston, Harris, Liberty, Matagorda, Montgomery, Walker, Waller, and Wharton counties (Figure 1). Over the 21-year period, Greater Houston added 2.7 million residents, growing by 59% from a total population of 4.3 million in 1997 to 6.8 million in 2017 (U.S. Census Bureau 2018). Greater Houston ranks as the fourth largest metropolitan area by population in the United States (Wilson et al. 2012). Urban centres are primarily restricted to Houston, Sugarland, and The Woodlands, which together house 88% of the region's total population (U.S. Census Bureau 2018). Outlying counties in the rural-urban interface consist largely of a network of interconnected towns and satellite communities surrounded primarily by agriculture, pasture, forest, and grassland.

2.2. Remotely-sensed data

All classifications were derived from Landsat satellite imagery spanning three satellite missions – the Landsat-5 Thematic Mapper (TM) for 1997–2011, the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) for 1999–2012, and the Landsat 8 Operational Land Imager (OLI) for 2013–2017 – and four Landsat World Reference System 2 (WRS-2) scenes: path/row 25/39, 25/40, 26/39, and 29/40 (Figure 1). All input imagery consists of radiometrically-calibrated and orthorectified Landsat Collection 1 Level-1 products conforming to prescribed criteria for < 10% cloud cover and possessing at least three phenological states per year: leaf-off (DOY 301–60), early growing season (61–180), and late growing season (DOY 181–300) (Appendix 1). Imagery was constrained to the calendar year of interest to ensure temporal precision in time series change detection,

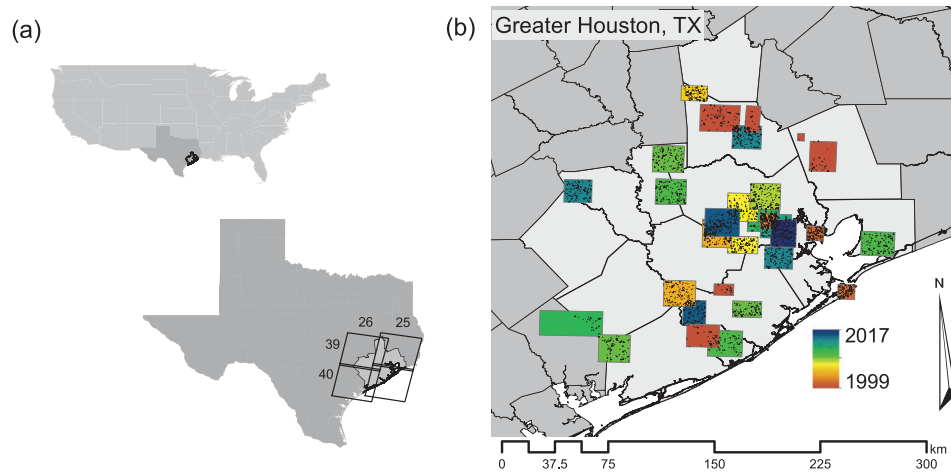


Figure 1. Greater Houston study area. (a) Study area extent (light grey) and four Landsat scene footprints with path/row designation (black outlines) superimposed on maps of the US and Texas; (b) County map with validation imagery extents (coloured by year) and validation samples (points).

and thereby precludes other commonly used predictor layers (e.g. DEMs) unavailable on an annual basis. Only pixels with high confidence in quality, as designated in corresponding Quality Assessment bands, were retained. For those years possessing sparse cloud-free imagery, lacking an acceptable range of acquisition dates, or otherwise heavily impacted by ETM+ Scan Line Corrector (SLC-off) data gaps, the cloud-cover and DOY criteria were relaxed. Given these constraints, a total of 262 Landsat scenes were used for the 21-year time series (Figure 2).

Training data for all classifications come from the U.S. Geological Survey (USGS) National Land Cover Database (NLCD) from 2001 (Homer et al. 2007), 2006 (Fry et al. 2011), and 2011 (Homer et al. 2015). Owing to trade-offs between classification accuracy and thematic precision, cover types were simplified to focus more acutely on urbanization trends (hereafter defined as land cover conversion from a non-Developed to a Developed class) rather than subtle ecological transitions such as wetland delineation, otherwise beyond the scope of the current study. Therefore, vegetation classes adopted from the NLCD's Anderson Level 2 typology were bifurcated into woody and non-woody vegetation whereby deciduous forest, evergreen forest, mixed forest, shrub/scrub, and woody wetlands were combined as 'Forest', while grassland/herbaceous, emergent herbaceous wetlands, and pasture/hay were merged as 'Grassland/Pasture'. All other classes occurring in the study area, as defined by the NLCD, were retained (Table 1). All Landsat images and NLCD classified maps were reprojected from their native coordinate system to a shared State Plane coordinate system, clipped to the 13-county study area, and buffered outward by 90m (~ 3 pixels) on all sides to mitigate against edge effects in spatial filtering.

Validation imagery consists of 30 fine-resolution images from the IKONOS, Quickbird, and Worldview-2 satellite sensors (©2018, DigitalGlobe; NextView License) and two

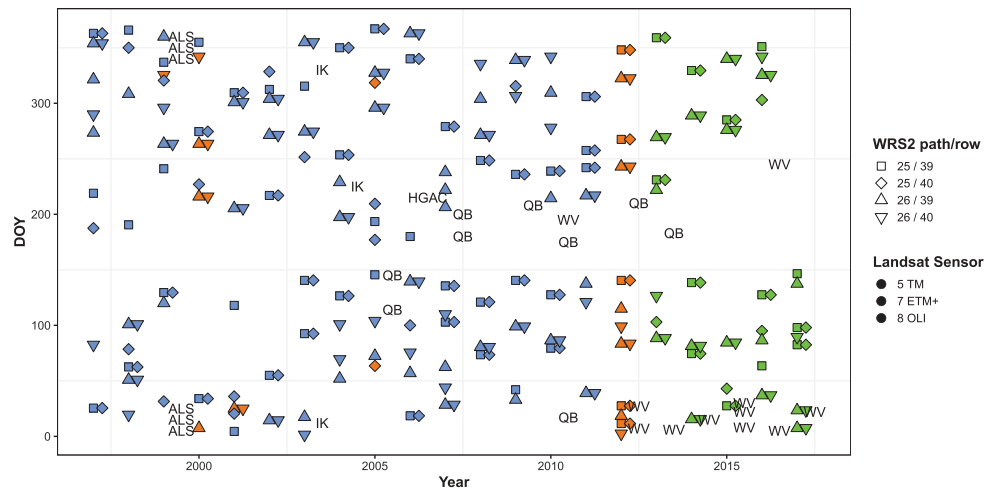


Figure 2. Distribution of Landsat scenes and fine resolution validation imagery. 262 Landsat scenes in total. Thirty validation images depicted by sensor abbreviation (ALS – Andrew Lonnie Sikes and HGA – Houston Galveston Area Council aerial imagery; as well as IK – IKONOS, QB Quickbird, and WV – Worldview-2 satellite imagery).

Table 1. Land cover class NLCD comparison.

Cover class	Corresponding NLCD class (code)
Barren/Sand	Barren Land – Rock/Sand/Clay (31)
Developed-Open	Developed, Open Space (21)
Developed-Low	Developed, Low Intensity (22)
Developed-Medium	Developed, Medium Intensity (23)
Developed-High	Developed, High Intensity (24)
Cultivated Crops	Cultivated Crops (82)
Grassland/Pasture	Grassland/Herbaceous (71)
	Emergent herbaceous wetlands (95)
	Pasture/Hay (81)
Forest	Deciduous Forest (41)
	Evergreen Forest (42)
	Mixed Forest (43)
	Shrub/Scrub (52)
	Woody Wetlands (90)
Water	Open Water (11)

airborne platforms: Andrew Lonnie Sikes and Houston Galveston Area Council aerial imagery (Kinder Institute 2018) (Figures 1(b) and 2; Appendix 2).

2.3. Class membership probabilities

Preliminary posterior class membership probabilities were derived from Landsat imagery based a three-step process: (1) image and band compositing using principal components analysis (PCA), (2) automatic adaptive signature generalization (AASG) (Gray and Song 2013; Dannenberg, Hakkenberg, and Song 2016), and (3) random forest (RF) supervised classification (Breiman 2001) (Figure 3). Prior to classification, all Landsat bands in a given image (6 bands, excluding thermal and fine-resolution panchromatic bands) were reduced to their first three PCA axes (PCA3) for computational efficiency. Concurrently, all images in a given year (i.e. 3–7 images times 6 bands per image = 18–

42 raw bands) were reduced to their first 10 PCA axes (PCA10), which represent > 99% of total variation in each annual image stack.

Next, to streamline the otherwise inconsistent and labour-intensive process of selecting training and predictor data in spatially-coincident multi-temporal image stacks, we employed the AASG algorithm. AASG first delineates stable (no-change) sites between images, defined as core areas within a scene whose cover class designation remains unchanged between the date of a reference image (I_R) and a target image (I_T). Stable sites are algorithmically determined by first selecting pixels within a pre-defined distance (c) from the mean (μ) of the image difference histogram (ΔI), where:

$$\Delta I = I_R[\cdot, 1] - I_T[\cdot, 1] \quad (1)$$

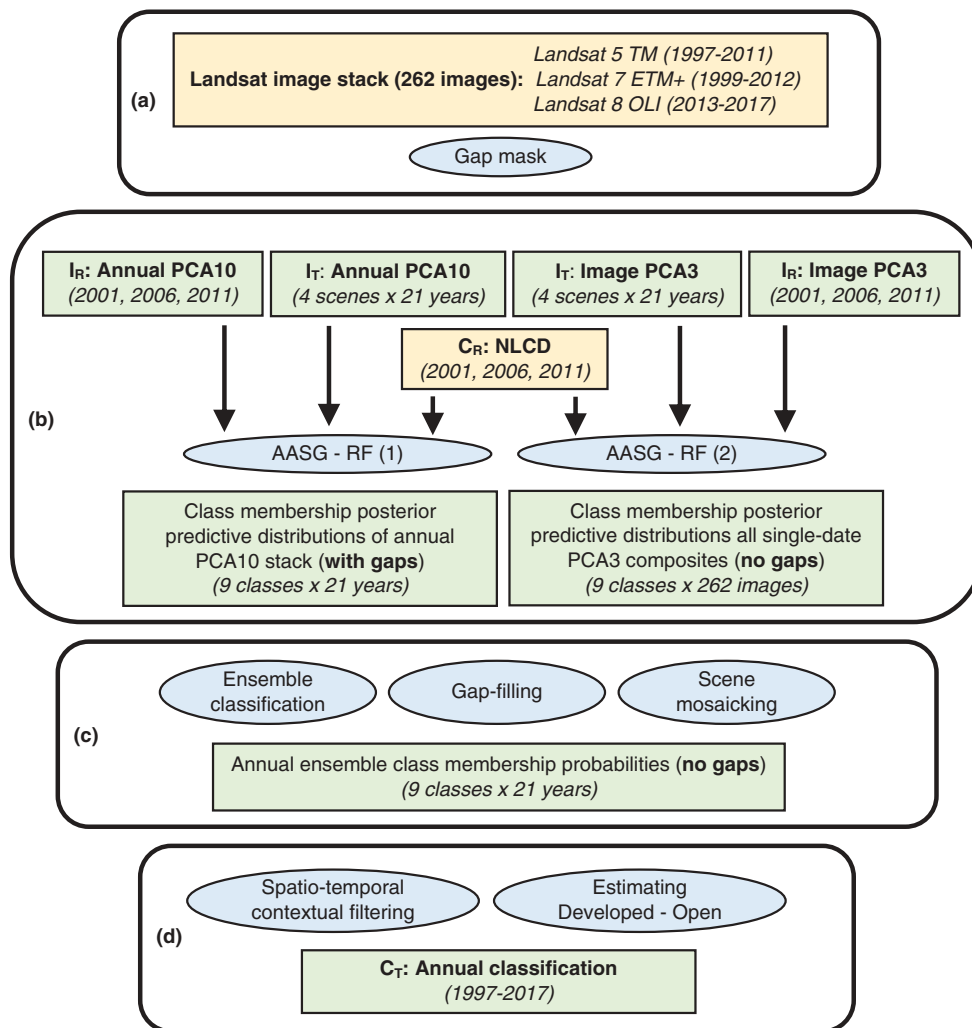


Figure 3. Methods flowchart. (a) Input imagery and cloud/shadow/SLC-off masking; (b) Model training and prediction, generating class membership posterior distributions; (c) Annual ensemble classification, including gap-filling and scene mosaicking; (d) Spatio-temporal filtering and derivation of final land cover time series. Inputs (yellow); process (blue); outputs (green).

such that, in this case, $[,1]$ corresponds to the first PCA axis derived from all spectral bands. Stable sites are selected from within the interval:

$$\mu_{\Delta I} \pm c_k \times \sigma_{\Delta I} \quad (2)$$

where $\mu_{\Delta I}$ and $\sigma_{\Delta I}$ is the mean and standard deviation of ΔI , respectively, and c_k is a class-specific threshold parameter for each class k . Candidate stable sites are additionally subjected to a class-specific spatial erode filter to mitigate against errors arising due to image misregistration and edge effects along class boundaries. Once delineated, scene-specific spectral signatures can be sampled from stable sites in both I_R and I_T images, and subsequently combined with a reference classification (C_R) corresponding to the date of the I_R for model training and prediction. By adapting to the unique atmospheric, radiometric, and phenological characteristics of each image, the AASG procedure facilitates automated image ingestion and classification processes that require neither atmospheric correction nor data normalization, while maintaining semantic consistency in class definitions between the reference and target classification (C_T) (Song et al. 2001).

As an automated training and predictor data selection algorithm, AASG is agnostic to the choice of classifier. We ultimately selected RF, an ensemble of classification trees based on votes across bootstrap replicates, for its computational efficiency and its record of high performance in terms of predictive accuracy and generalizability (Belgiu and Drăgu 2016). The nonparametric RF algorithm produces highly accurate and unbiased predictions that efficiently handles highly collinear neighbouring predictor pixels in each stable site, is robust to noise, and largely immune to over-fitting – of interest due to the requirement that identical training data generalize to so many different target images in the Landsat stack (Gislason, Benediktsson, and Sveinsson 2006).

RF classification models for each scene/year were parameterized with 200 trees per model, with 3 predictors sampled at each split using training data from AASG-defined stable sites (Maxwell, Warner, and Fang 2018). Each training class was proportional to the relative abundance of each reference class and capped at 100,000 pixels per class (Chen, Liaw, and Breiman 2004). The three NLCD reference classifications (C_{R-2001} , C_{R-2006} , and C_{R-2011}) were paired with reference imagery from each respective year (I_{R-2001} , I_{R-2006} , and I_{R-2011}), and applied to the most temporally-proximate target imagery for all 21 years (i.e. I_{R-2001} and C_{R-2001} corresponds with $I_{T-1997} - I_{T-2003}$, while I_{R-2006} and C_{R-2006} was paired with $I_{T-2004} - I_{T-2008}$, and I_{R-2011} and C_{R-2011} with $I_{T-2009} - I_{T-2017}$). Raw predictions, in the form of posterior membership probabilities (p) for each class (i), are based on the distribution of ‘votes’ from the ensemble of classification trees in the RF classifier, such that:

$$\sum_{i=1}^k p_i = 1 \quad (3)$$

for k classes per pixel (Wang et al. 2015). All RF models were run using the randomForest package (Liaw and Wiener 2002) and derived products and analyses were calculated using the raster package (Hijmans 2017) in the software R, v. 3.3.1 (R Core Team 2017).

2.4. Annual ensemble classification

Annual PCA10 composites (see [Section 2.3](#)), which incorporate spectral data from multiple images across three phenological states within a given calendar year, serve as the primary predictor in all classifications ([Figure 4\(a\)](#)). Owing to data gaps in the PCA10 predictor set – which represent the superset of all algorithmically delineated clouds and cloud shadows (Zhu, Wang, and Woodcock 2015) as well as ETM+ SLC-off gaps and radiometrically-saturated or contaminated pixels identified in quality assessment bands ([Figure 4\(b\)](#)) – a parallel classification was simultaneously conducted on the PCA3 composite from each single image. Specifically, AASG-RF was implemented on each PCA3 in the annual stack and used to generate per-pixel posterior predictive distributions for each class ([Figure 4\(c-d\)](#)). From these posteriors, an ensemble prediction was derived from the geometric mean of the set of 3–7 posterior classification probabilities in a given year and used as the basis for designating pixels' class membership ([Figure 4\(e\)](#)). These classified pixels were then used to fill data gaps in the original PCA10 classification ([Figure 4\(f\)](#)). Unlike gap-filling algorithms that interpolate pixel values before classification, this two-part classification procedure ensures all classifications are derived from original reflectance values, thereby retaining pixel-level spatial consistency (Yin et al. 2017). This ensemble classification approach utilizes the added information content of the full stack of all acceptable imagery in a calendar year to mitigate the potential for contagion or classification error of any one image, as well as inter-image pixel misalignment due to discrepancies in georegistration. Because PCA3s are only

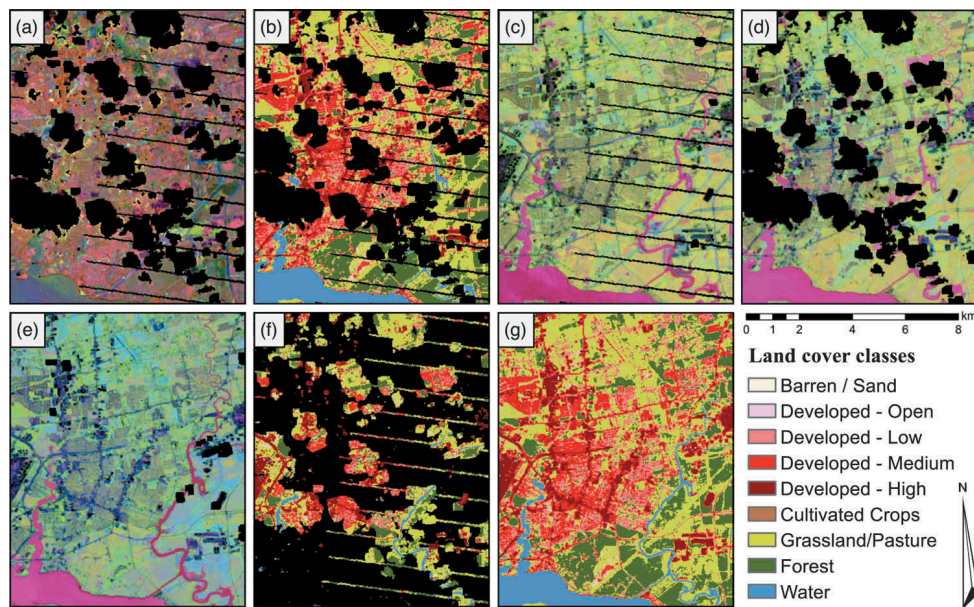


Figure 4. Multi-date classification procedure. (a) Annual PCA10, with ETM+ SLC-off and cloud/shadows masked (black); (b) PCA10 classification with data gaps (black); (c-e) single-date PCA3 image composites with data gaps (black); (f) classification of PCA10 gaps based on annually-aggregated, mean membership probabilities of all PCA3 classifications; (g) gap-filled classification (combining panels b and e). Bounding box corresponds with [Figure 6\(a\)](#), box 2.

impacted by data gaps resulting from stochastic phenomena (e.g. cloud location) in any one image, overlap in missing data pixels for all multi-temporal images in a given year is extremely rare, and can be interpolated during temporal stabilization (See Section 2.4.1). To ensure a seamless transition between neighbouring scenes, mean classification probabilities in the 2–4 overlapping scene edge areas were used to replace those produced for each scene.

2.5. Spatio-temporal filtering

2.5.1. Spatial-temporal contextual filtering

To mitigate against error propagation due to misclassification and ensure consistency in automated time series classifications, we adopted a spatio-temporal contextual filtering approach that exploits two statistical properties of the classified time series – namely, spatial autocorrelation and temporal dependence – to identify potential spurious classifications and adjust them accordingly (Lu and Weng 2007; Li et al. 2014). Contextual filters exploit information between a target pixel and neighbouring pixels within spatial and temporal windows of varying size to impose constraints on the final classification of the target pixel. Contextual filtering consisted of three steps: (1) temporal smoothing, (2) spatial filtering, and (3) label modification for illogical temporal transitions.

For temporal stabilization of classification probabilities, especially where class probabilities exhibit pronounced peaks and troughs in the temporal domain, we applied a temporal low pass filter using a Gaussian kernel in a five year window (Hamilton 2015). Spatially-weighted kernel filters were then applied to each classification in the time series to remove spurious spatial heterogeneity (e.g. ‘salt-and-pepper’) in otherwise homogeneous land cover patches. In addition to spatial kernel filters, a minimum mapping unit (MMU) criteria was applied following Homer et al. (2015), whereby a 5-pixel MMU was required for all classes except Cultivated Crops (which required a 12-pixel MMU) and Developed classes, which were not subjected to the MMU requirement. Lastly, a rule-based label adjustment procedure was used to eliminate illogical temporal transitions in the time series identified when the class of maximum posterior probability exhibits pronounced fluctuations within a short time period (Wang et al. 2015; Zhang and Weng 2016). For example, for cover classes exhibiting relatively discrete spatial boundaries (e.g. the four Developed classes), a three-year temporal window ($t-1$, t , $t+1$) was employed such that the classification at time t was modified to that for time $t-1$, when $t-1 = t+1$ and $t \neq t-1$ (Pouliot et al. 2014; He, Lee, and Warner 2017). For land cover classes exhibiting more continuous temporal variation in land surface properties (e.g. Grassland/Pasture and Forest) a more conservative five-year temporal filter ($t-2$: $t+2$) was employed to distinguish long-term (genuine) trends from short-term (spurious) fluctuations (Cai et al. 2014).

2.5.2. Special consideration for the Developed-Open class

Following NLCD definitions, the four Developed classes – Open, Low, Medium, and High Intensity – are defined by impervious surface fractional covers of 0–20%, 20–49%, 50–79%, and 80–100%, respectively (Appendix 3). Of particular concern for the current study is the characterization of Developed-Open pixels that, being defined as < 20% impervious cover, would otherwise possess the spectral characteristics of the

predominant fractional cover class, such as water or vegetation. Because the Developed-Open class is defined by an impervious fractional cover far below 50%, its delineation in the NLCD protocol requires additional non-spectral data unavailable at annual time scales, as well as manual boundary delineation (Jon Dewitz, personal communication, 24 January 2018). And while the results of this resource-intensive process are highly satisfactory, the approach is neither reproducible nor feasible for automated classification at an annual scale. In response, several studies have simply eschewed classifying the Developed – Open class altogether (Sexton et al. 2013; Dannenberg, Song, and Hakkenberg 2018). However, as a central component of the low density, sprawling development characteristic of Greater Houston, as well as its disproportionate impact on urban flood risk, mitigation, and planning, we deemed it necessary to include a spectrally-determined, high-fidelity proxy for the Developed – Open class (Brody, Kim, and Gunn 2013).

We therefore approximated the Developed – Open class as all pixels falling within a aggregated urban extent, that otherwise do not possess the fractional impervious cover proportions defining the three higher-intensity – Low, Medium, and High – Developed classes. To do this, posterior probabilities from raw classifications were assessed to identify pixels whose RF modal vote prediction falls within one of the four Developed classes (Figure 5(b-c)). Because this spectrally-determined impervious layer fails to capture the full extent of the NLCD's Developed classes (including the partly manually-determined < 20% impervious Developed – Open) especially where vegetated yards or overhanging tree canopies in suburban areas were misclassified as vegetation, we applied a 3×3 and 5×5 anisotropic spatial filter to the classified output to identify interstitial and edge pixels that should be included within the urban base map (Figure 5(d)). Given this urban extent, the three higher-intensity – Low, Medium, and High Intensity – impervious classes (Figure 5(e)) were superimposed within the urban extent (Figure 5(f)), such that all remaining pixels are classified as Developed – Open. Approximated urban extents show a strong resemblance to concurrent NLCD maps (Figure 5(g)) with F-scores, representing the harmonic mean of the user's and producer's accuracies between the two urban extents, achieving values of 0.894, 0.887, 0.885 for 2001, 2006 and 2011, respectively (Appendix 4). Thereafter, the urban mask was updated annually in a manner consistent with other temporal filtering processes in addition to one illogical transition criterion based on an irreversibility assumption adopted from Gao et al. (2012): once a pixel is classified as one of the Developed categories for a minimum of three consecutive years, it is sufficiently unlikely to be converted again in the study time period. Supporting this assumption in other studies, there were zero pixels that transitioned from a Developed to a non-Developed category in the NLCD map of the study area from 2001 to 2011 (Homer et al. 2007, 2015) – a result similarly observed in Washington DC by Sexton et al. (2013).

2.6. Accuracy Assessment

Maps were assessed for accuracy by comparison with coincident NLCD maps and via a three-part validation procedure with independent, multi-temporal, fine resolution imagery based on a sampling and response design modified from Olofsson et al. (2014): (1) full class, crisp accuracy assessment; (2) reduced class, crisp accuracy assessment; and (3) fuzzy

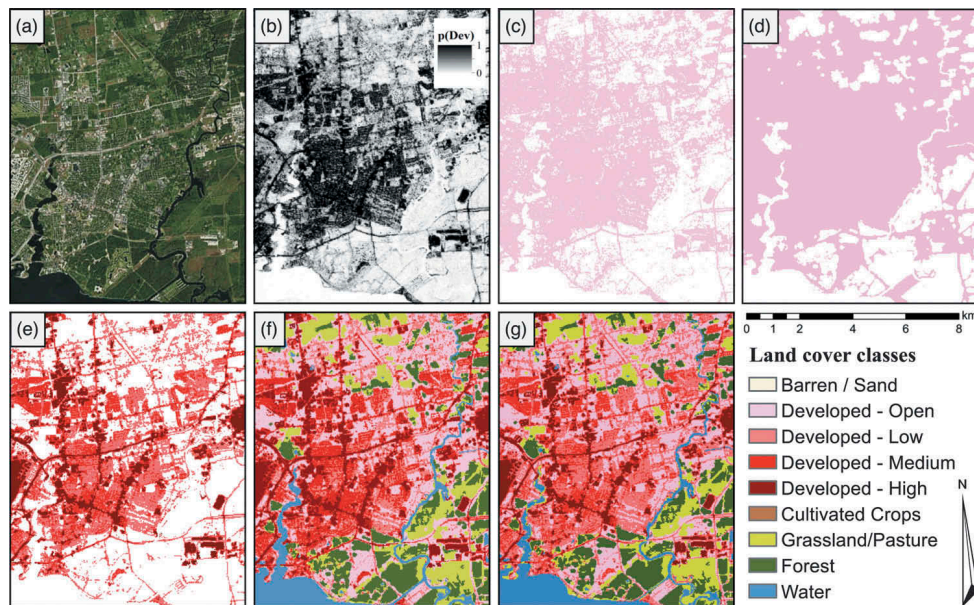


Figure 5. Estimation of urban extent. (a) Fine resolution aerial reference imagery; (b) RF posterior probability of combined Developed classes from 2012; (c) raw urban extent derived from modal posterior probabilities; (d) spatially-filtered urban extent; (e) classified pixels in the Developed – Low, Medium, and High classes; (f) final classification; (g) coincident and concurrent NLCD classification. Bounding box corresponds with Figure 6(a), box 2.

accuracy assessment. First, classified maps were compared with spatially, temporally, and thematically coincident NLCD maps to assess overall agreement (O_{AG}), producer's agreement (P_{AG}), and user's agreement (U_{AG}) for the three nominal years where the two products overlap (e.g. 2001, 2006, and 2011) (Congalton 1991).

Second, we conducted a multi-temporal independent accuracy assessment using a stratified random sampling design whereby samples corresponding to the nominal resolution of classified maps were established in advance in 30 fine-resolution ($\leq 3\text{m}$) satellite and aerial images (Appendix 2). Validation imagery is adequately distributed in space (as measured by correspondence in total areal cover by class in the full study area extent versus that for reference imagery only) and time (14 of 21 years) throughout the study area (Figure 1(b); Appendix 5). Total sample size ($n = 3036$) across the 14 reference dates was determined by a priori expectations for the average standard error in the overall agreement with the three NLCD products, and adjusted upwards to account for rare classes (Olofsson et al. 2014). All samples were allocated proportionally by cover class strata and across reference imagery by year and spatial extent. Specific sample locations were determined independently from AASG stable sites and, given stratification constraints, randomized.

Thereafter, trained technicians conducted a blind interpretation of land cover within the areal extent of each sample pixel, allowing for mixed pixels and other ambiguities by assigning proportional membership when class identity was not otherwise unambiguous (e.g. membership score $p_k \neq 1$). To ensure a monotonic ranking, no two membership probabilities were equal. Due to the possibility of interpretation error and inconsistency,

all samples were classified by more than one technician, with disagreement in the class of maximum probability leading to secondary expert review. Thereupon, accuracy assessment results follow standard protocols for reporting overall accuracy (OA), user's accuracies (UA), and producer's accuracies (PA), with all corresponding 95% confidence intervals (CIs) based on the area-weighted population error matrix (Olofsson et al. 2014; Foody 2002). As a single statistic for classification accuracy, the area-weighted overall accuracy was favoured to alternative approaches like the kappa coefficient (Pontius and Millones 2011).

Owing to the relatively coarse resolution of Landsat imagery in relation to end-member fractional cover, as well the inherent subjectivities in assigning a single 'crisp' class membership in reference imagery, crisp accuracy may have limited utility, especially for highly heterogeneous urban land cover (Foody 2002). The uncertainty and ambiguity inherent in crisp accuracy assessments is non-trivial and especially apparent in Developed mixed pixels which, despite existing on a continuum of surface imperviousness, are binned into discrete categories. Added to this uncertainty is the lack of confidence in the consistency of reference labels based on technicians' visual estimate of surface imperviousness. We therefore implemented a fuzzy accuracy assessment based on a translation of visually determined membership probabilities, using a three-level linguistic-measurement scale to characterize the magnitude of membership probability, with the highest single class probability defined as absolutely right ('Right'), the second highest as a good answer ('Good'), and all other non-zero probabilities (maximum of two) assigned as reasonable or acceptable ('Acceptable') (Woodcock and Gopal 2000; Foody 2002). Because the 'Right' category is, by definition, equivalent to crisp UAs in the area-weighted population error matrix, we limit results to the 'Good' and 'Acceptable' categories.

2.7. Cover class area estimation

Annual class area estimates were determined based on a stratified estimator of areal proportions derived from independent reference imagery. Accepting that the accuracy assessment sampling design yielded estimates with relatively small standard errors, as well as the premise that the quality of the independent reference imagery is superior to that of the map classification, class areas can be estimated by multiplying area proportions derived from the population error matrix of the independent reference imagery (i.e. column totals of the contingency table) by the total map area (Stehman 2013). This sampling design likewise allows for the estimation of unbiased standard errors for each class area (Olofsson et al. 2014). For simplicity, the contingency table used for area estimates was constrained to single, crisp membership consisting of the highest probability class among all independent samples. While the derivation of area estimation parameters from an error matrix populated with crisp set memberships tentatively assumes mutually exclusive and collectively exhaustive categories at odds with fuzzy logic, it simultaneously allows class areas to sum to one, and thereby better facilitates consistent inter-annual comparisons of class areas (Woodcock and Gopal 2000).

3. Results

3.1. Annual land cover time series

Classified maps for the 21-year time series (Figure 6) show strong visual fidelity to known land cover patterns and demonstrate the expansive scale of the Greater Houston region, which in the absence of significant topographic constraints assumes a symmetrical, hub and spoke urban form (Galster et al. 2001). As such, developed areas are tightly clustered in the urban core, while sprawling suburbs expand outwards in all directions along major transportation corridors and emerging satellite communities populate the urban periphery where they have leap-frogged non-Developed classes (Jaret et al. 2009). The outer periphery consists largely of Cultivated Crops, Grassland/Pasture, and Forest cover types, within which older ranching and agricultural settlements and communities are scattered. Zoomed subsets of the study area demonstrate the capacity for characterizing the texture of intergrading impervious surfaces across an urban density gradient (Figure 7).

3.2. Classification accuracies

While differences exist between this land cover time series and coincident NLCD maps (e.g. thematic categories), they nonetheless still demonstrate a substantial degree of agreement (Table 2). The largest disparities between the two products occur with Barren/Sand and the four Developed classes, while natural and semi-natural classes (e.g. Forest, Grassland/Pasture, and Water) exhibit close agreement for concurrent dates, on the order of 73–99%. Based on a random stratified sampling design with multi-date fine-resolution images, we found the full

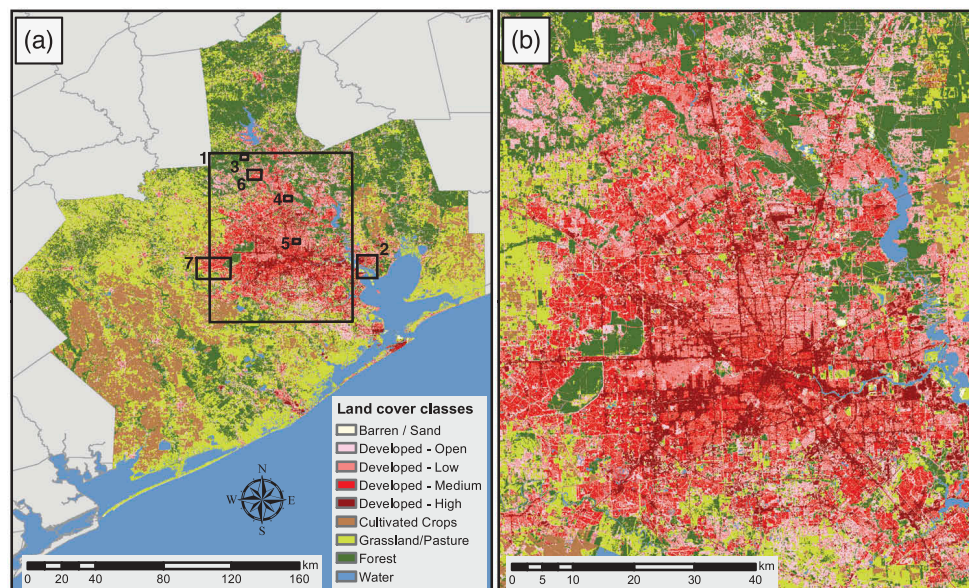


Figure 6. Land cover classifications of Greater Houston (2017). (a) HGA study area; (b) Houston (box 1). Bounding box 2 (Figures 4 and 5); boxes 3–5 (Figure 7); boxes 6–7 (Figure 11).

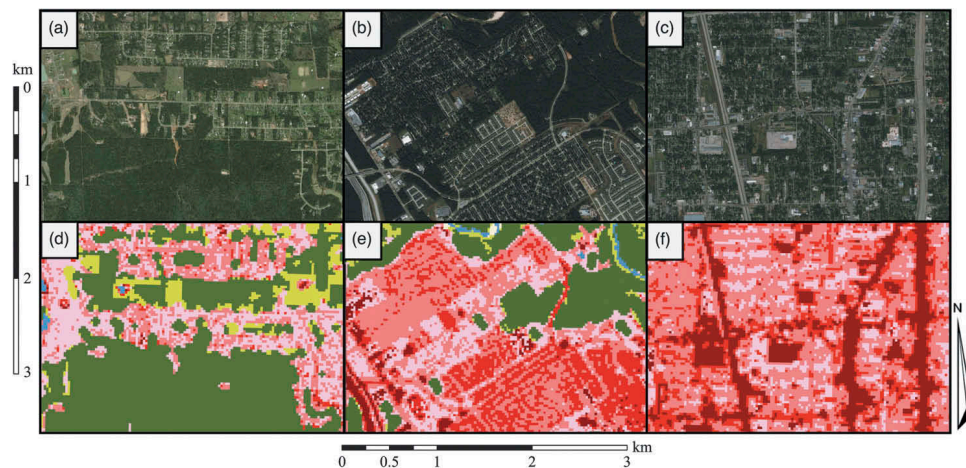


Figure 7. Zoomed urban classification insets. Fine resolution aerial reference imagery (a–c) and corresponding classifications (d–f) corresponding to 2017, with increasing levels of urbanization (from light to dark red). Specific locations correspond to bounding boxes in [Figure 6\(a\)](#): (a, d) box 3; (b, e) box 4; (c, f) box 5. Classification coloration is consistent with legends in [Figure 4–6](#).

nine-class maps to achieve an overall accuracy of 78% ($\pm 1.5\%$), with user's accuracies lowest for the Developed classes, mostly due to confusion among the different intensities of Developed land rather than confusion with non-Developed cover types ([Table 3](#); Appendix 6). Accordingly, with all Developed classes merged, overall accuracy reaches 86% ($\pm 1.4\%$) ([Table 4](#)). Fuzzy accuracy assessment results demonstrate a 90.6% ($\pm 1.5\%$) accuracy for 'good' matches and 94.2% ($\pm 1\%$) accuracy for 'acceptable' agreement ([Table 5](#)).

3.3. Greater Houston land cover change area estimates

Unbiased cover class areas were estimated from areal proportions in the the population error matrix ([Table 3](#)). The largest land cover changes observed in the study area occurred in the Developed classes, especially the Developed – Medium category, which grew by 62% over the 21-year period (2.3% compound annual) and the Developed – High class, with 52% total growth (2.0% compound annual) observed ([Figure 8](#)). In total, combined Developed classes witnessed an increase of

Table 2. Agreement with NLCD maps for 2001, 2006, and 2011 (%). U_{AG} – user's agreement; P_{AG} – producer's agreement; O_{AG} – overall agreement. NLCD as reference.

	2001		2006		2011	
	U_{AG} (%)	P_{AG} (%)	U_{AG} (%)	P_{AG} (%)	U_{AG} (%)	P_{AG} (%)
Barren/Sand	71.0	30.5	63.9	28.9	67.6	29.6
Developed-Open	58.1	76.9	53.9	74.9	52.0	74.4
Developed-Low	48.1	43.6	44.0	48.0	43.2	49.4
Developed-Medium	60.7	45.6	57.3	51.4	60.0	50.7
Developed-High	58.0	71.0	54.7	73.7	56.4	73.1
Cultivated Crops	80.0	76.7	80.1	76.3	80.1	76.2
Grassland/Pasture	75.3	81.5	75.8	79.9	76.1	78.9
Forest	86.8	76.9	88.2	75.1	87.1	75.2
Water	84.4	99.5	84.9	99.4	85.7	97.7
O_{AG}	75.7		74.9		74.3	

Table 3. Area-weighted confusion matrix (full). *UA* – user’s accuracy; *PA* – producer’s accuracy; *OA* – overall accuracy. Accuracies are listed as proportions of the total study area, followed by 95% confidence intervals.

		Reference									
		Barren/Sand (%)	Developed -Open (%)	Developed -Low (%)	Developed -Med (%)	Developed -High (%)	Cultivated Crops (%)	Grassland/Pasture (%)	Forest (%)	Water (%)	UA (%)
Map	Barren/Sand	0.2	0	0	0	0	0	0	0	0	96.1 ± 4.4
	Developed-Open	0	3.6	2.9	0.7	0.3	0	0.5	0.3	0	42.2 ± 5.6
	Developed-Low	0	0.4	2.7	1.4	0.3	0	0	0	0	54.8 ± 5.6
	Developed-Med	0	0.1	0.6	1.7	0.6	0	0	0	0	56.2 ± 5.1
	Developed-High	0	0	0.1	0.4	1.7	0	0	0	0	76.5 ± 4.6
	Cultivated Crops	0	0.1	0	0	0	8.2	3.0	0.3	0	69.2 ± 7.2
	Grassland/Pasture	0	1.7	0.5	0	0	1.3	29.5	1.1	0.2	85.7 ± 2.5
	Forest	0	1.2	0.4	0	0	0	1.7	20.6	0.2	85.4 ± 3.0
	Water	0.1	0	0	0	0	0	0.5	0.1	9.7	90.7 ± 3.9
	PA	45.8 ± 11.4	50.1 ± 33.4	36.4 ± 23.8	40.2 ± 21.6	56.9 ± 14.9	85.6 ± 12.8	83.3 ± 4.1	91.9 ± 2.6	94.7 ± 2.1	OA 78.0 ± 1.5

Table 4. Area-weighted confusion matrix (reduced). Developed classes combined. *UA* – user's accuracy; *PA* – producer's accuracy; *OA* – overall accuracy. Accuracies are listed as proportions of the total study area, followed by 95% confidence intervals.

		Reference						
		Barren/ Sand (%)	Developed -combined (%)	Cultivated Crops (%)	Grassland/ Pasture (%)	Forest (%)	Water (%)	<i>UA</i> (%)
Map	Barren/Sand	0.2	0	0	0	0	0	96.1 ± 4.4
	Developed-combined	0	18.0	0	0.4	0.1	0	96.2 ± 1.0
	Cultivated Crops	0	0.2	8.2	3.0	0.3	0	69.2 ± 7.2
	Grassland/Pasture	0	2.3	1.3	29.5	1.1	0.002	85.7 ± 2.5
	Forest	0	1.6	0	1.7	20.6	0.002	85.4 ± 3.0
	Water	0.1	0.2	0	0.5	0.1	0.097	90.7 ± 4.5
	<i>PA</i>	45.2 ± 11.4	80.8 ± 10.6	85.8 ± 12.7	83.9 ± 4.0	92.4 ± 2.4	94.9 ± 2.0	<i>OA</i>
								86.2 ± 1.4

Table 5. Fuzzy accuracy assessment. *UA* – user's accuracy; *OA* – overall accuracy, followed by 95% confidence intervals. Fuzzy linguistic scale following Woodcock and Gopal (2000): good answer ('Good') and reasonable or acceptable ('Acceptable').

		Reference	
		<i>UA</i> 'Good' (%)	<i>UA</i> 'Acceptable' (%)
Map	Barren/Sand	96.3 ± 4.2	97.5 ± 3.4
	Developed-Open	71.8 ± 5.6	86.5 ± 4.4
	Developed-Low	83.0 ± 3.8	97.2 ± 1.7
	Developed-Med	83.8 ± 3.8	95.3 ± 2.1
	Developed-High	92.5 ± 2.9	96.2 ± 2.0
	Cultivated Crops	85.9 ± 5.5	87.8 ± 5.1
	Grassland/Pasture	95.1 ± 1.6	96.4 ± 1.3
	Forest	93.0 ± 2.2	95.2 ± 1.8
	Water	95.8 ± 2.7	95.8 ± 2.7
	<i>OA</i>	90.6 ± 1.1	94.2 ± 1.0

2040 km² ± 400 km² (Figure 9), with the Low, Medium, and High Intensity Developed classes accounting for 41%, 34%, and 21% of that growth, respectively. The remaining 4% of the change is attributable to expansion of the Developed – Open class. While the higher-intensity Developed classes experienced the largest rates of change over the 21-year period, developed cover in the study area was still dominated by the low-intensity, spatially-dispersed urban morphology of the Developed – Open (33% of total developed area) and Developed – Low (34% of total developed area) categories. Growth in Developed classes was largely offset by declines of –4.3% and –15.6% (–0.2% and –0.8% compound annual) in the Grassland/Pasture and Forest classes, respectively. In total, Forest cover decreased by 1350 km² (± 460 km²), while Cultivated Crops and Grassland/Pasture experienced nonsignificant declines of 100 km² (± 490 km²) and 550 km² (± 670 km²), respectively (Figure 9).

4. Discussion

4.1. Land cover change trends in Greater Houston

Areal change maps corroborate what census data likewise indicate: rapid growth in the 13-county region over the past two decades, whereby an estimated 59% growth in population corresponds to a 30.3% (± 3.3%) increase in Developed cover (U.S. Census

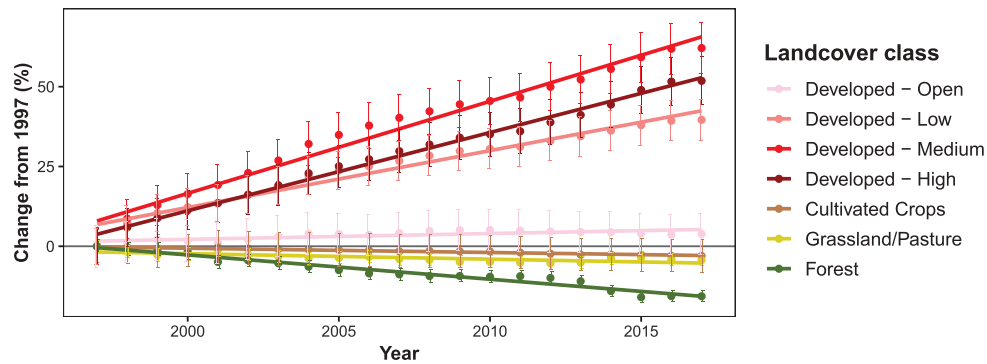


Figure 8. Greater Houston land cover change rates. Barren/sand and water are excluded. Linear regression line added for reference. Error bars based on standard error.

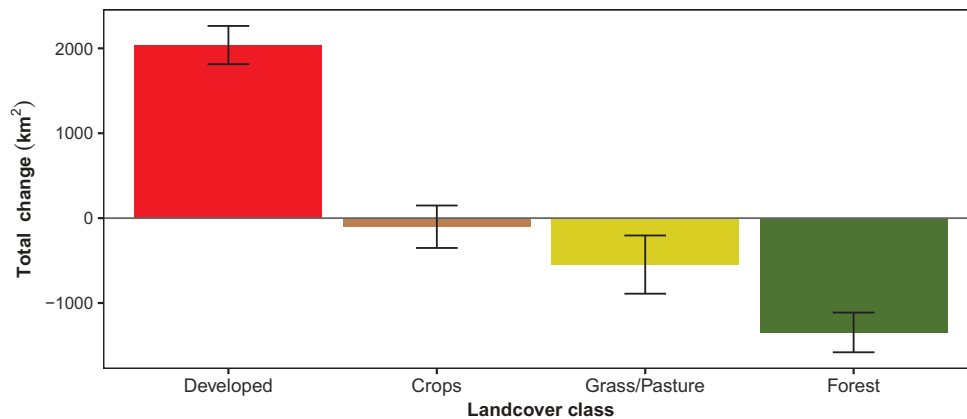


Figure 9. Greater Houston land cover change totals between 1997–2017. Total change and standard error for the four largest land cover change classes.

Bureau 2018). Despite Houston's ranking as among the most sprawling large American cities as measured by density and nuclearity (Galster et al. 2001), that the rate of urbanization is half that of population growth reflects some degree of densification, however modest. Notably, the largest proportional land cover changes observed in the study area occurred in the higher density Developed classes, especially the Developed–Medium and Developed–High categories, which grew by 62.1% ($\pm 9.8\%$) and 51.8% ($\pm 9.4\%$), respectively, versus the lower density Developed–Open and Developed–Low categories, which grew by 38.8% ($\pm 9.1\%$) and 39.6% ($\pm 8.5\%$), respectively. This finding of increased high-density growth largely corroborates the conclusions of other studies which find land availability constraints and growing commute times in Houston to be primary factors driving infill and multi-story developments (Brody, Kim, and Gunn 2013).

Increases in Developed cover were mostly offset by 4.3% ($\pm 2.6\%$) and 15.6% ($\pm 2.7\%$) declines in the Grassland/Pasture and Forest categories, respectively. The disparity in the magnitude of the positive versus negative change rates in the zero-sum game of land

cover conversion is explained by considering the vastly larger total area of the Grassland/Pasture and Forest classes in the study area versus the Developed categories. The relatively smaller decline in the Grassland/Pasture category compared with Forest cover is due to the far greater frequency of Forest conversion to Grassland/Pasture (e.g. deforestation) compared with the reverse trend (e.g. afforestation/reforestation). Based on a comparison with the National Oceanic and Atmospheric Administration's Coastal Change Analysis Program's land cover product, 14% of all land cover urbanized between 1997–2017 was classified as wetlands prior to conversion (NOAA C-CAP 2011). The magnitude of wetland conversion over the past two decades in the Greater Houston area has important implications for wetlands ecological conservation, storm water management, and flood hazards (Bullock and Acreman 2003).

While bi-temporal change detection provides an estimate of net land cover change with associated uncertainties, it is insufficient for characterizing spatially-variant change trajectories as well as temporal dynamics of urbanization morphologies (Yu and Zhou 2017). Multi-temporal classifications, on the other hand, are capable of detecting higher-order dynamics of land cover change (Li, Gong, and Liang 2015; Song et al. 2016). Periodic fluctuations in land cover growth trajectories are evident in the land cover time series, with the timing and magnitude of acceleration in the growth of urbanization mirroring the periodicity observed in socio-economic indicators including total population, Greater Houston's Gross Domestic Product (GDP), and the Harris County Housing Price Index (HPI) (Figure 10). Over the 21-year period, the rate of urbanization peaked between 2005–2007, followed by a considerable reduction relative to baseline growth after the start of the 'Great Recession' in the United States in late 2007. Interestingly, the timing of satellite-observable development is temporally offset from the underlying socio-economic forces partly driving it, shedding light on the magnitude of the temporal lag between the two related trends.

The spatial imprint of temporal processes of urbanization is particularly visible in change year maps. Using the example of The Woodlands and Cinco Rancho large-scale developments, we observe that while both exhibit some similar growth characteristics (e.g. expansion from an initial seed area), their growth morphologies are in fact quite different (Figure 11). For example, the stringently-zoned western extension of The Woodlands expands within a constrained area bounded by green space to the north, west, and south. The largely unzoned Cinco Rancho, on the other hand, expands outward in all directions, largely undeterred by zoning, topography, or hydrology in the process of converting former agricultural land to large-scale residential developments (Qian 2010). These individual developments exemplify the scale and pace of urbanization in the Greater Houston area, with the former area adding 27 km² (6.6% compound annual) in Developed cover over the 21-year period, while the latter added 115 km² (5.7% compound annual).

4.2. Classification accuracy

While no one statistic is singularly authoritative in validating the accuracy of dense land cover time series, the use of multiple assessments helps to clarify spurious or misleading confusion in the crisp classifications, while simultaneously providing a more robust ceiling for actual (rather than sampled) map accuracy. Among all classes, crisp

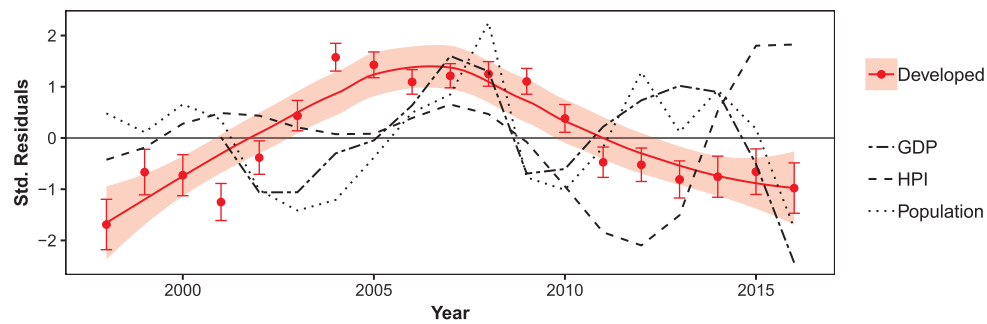


Figure 10. Land cover and socio-economic trends in the HGA. Standardized residuals from the slope (β_1) of a linear regression of urbanization, Greater Houston's Gross Domestic Product (GDP), Harris County House Price Index (HPI), and population. Land cover points and standard error bars represent class-specific areal estimates. The land cover trend line is represented by a loess function, plus 95% confidence interval. Socio-economic data from U.S. Bureau of Economic Analysis (2018) and U.S. Census Bureau (2018).

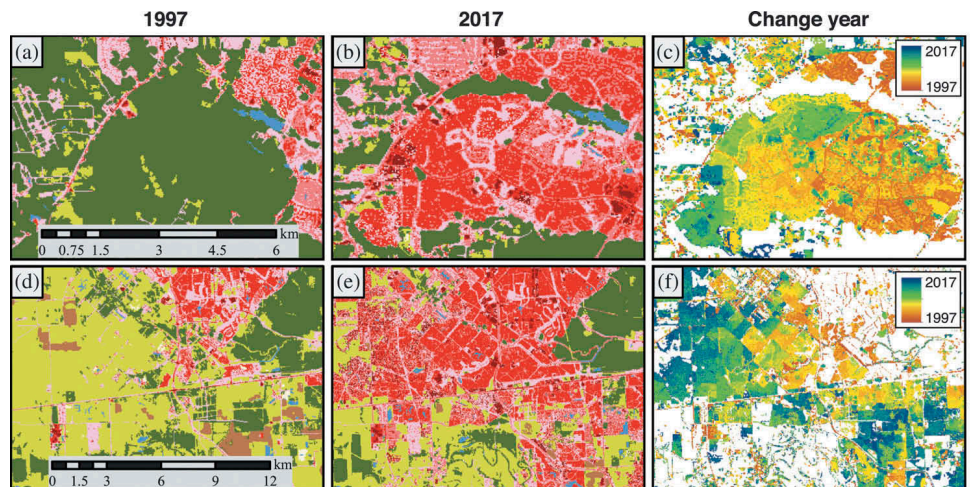


Figure 11. Change year maps for large-scale developments. The Woodlands, corresponding with bounding box 6 in Figure 6(a) (a-c) and Cinco Ranch, corresponding with bounding box 7 in Figure 6(a) (d-f). Classification coloration is consistent with legends in Figure 4–6, with darker reds indicating higher proportions of impervious surface.

classifications of Developed cover exhibited the lowest per-class accuracies owing largely to the subjectivity inherent in technicians' assignment of a single imperviousness value to the spatially complex, multi-endmember impervious cover types (Wang, Huang, and De Colstoun 2017; Weng 2012). Per-class agreement with the NLCD was likewise relatively low for these four Developed classes, though when combined into a single Developed class, user's accuracies achieve 96% overall accuracy and 89% agreement with the NLCD. It should be stressed that inference of accuracy from a test of agreement is problematic owing to the lack of an unambiguous reference map (errors exist in both products). Fuzzy accuracy assessments largely compensate for these misleadingly low

accuracies for the four distinct Developed classes, albeit at the expense of thematic precision, reaching 90–94% overall accuracy (Woodcock and Gopal 2000).

The Cultivated crops and Grassland/Pasture classes exhibited significant confusion, partly owing to the ambiguity regarding the taxonomic identity of vegetation in the two classes, as well as uncertainty in labelling samples in reference imagery. Confusion was likewise observed in the Barren/Sand and Cultivated Crop classes, which both tend to exhibit high reflectance values that may be easily confused with impervious surfaces (Wang, Huang, and De Colstoun 2017; Wickham et al. 2017). Furthermore, Barren/Sand (< 1% of total area) may represent a transitional state in the urbanization process (ground clearing and early construction) that could simultaneously be accurately characterized as a Developed class.

The accuracy of this Greater Houston land cover product generally compares favourably to those observed in similar studies, though caution is advised with direct comparison owing to idiosyncrasies in ground cover complexity among regions, as well as the distinct differences in spatial and thematic resolution, reference data, and assessment method (Gómez, White, and Wulder 2016). In a meta-analysis of over 500 studies between 1989 and 2003, Wilkinson (2005) observed a mean accuracy of 76% (15.6% sd). Furthermore, Herold et al. (2016) notes that map accuracies since 2011 generally range from 61% to 87%. Interestingly, despite the advancements in satellite data acquisition and classification algorithms, classification accuracies have not improved significantly in the past 30 years (Herold et al. 2016; Yu et al. 2014).

4.3. Towards fine-resolution, large-extent, annual land cover time series

The demand for map products capable of assessing increasingly fine-scale spatio-temporal dynamics over large extents and long durations has accelerated in recent years for research fields spanning the realms of urban socio-economics, hazard and risk mitigation/reduction, and ecosystem modelling (Jensen and Cowen 1999; Yu and Zhou 2017). To meet this demand, international efforts have proceeded swiftly to operationalize continuous, wall-to-wall monitoring of land cover change across the globe. The fast pace of satellite deployments over the past few years, coupled with the profusion of increasingly sophisticated data fusion techniques, has enabled near-daily monitoring of the Earth surface (Zhu et al. 2015; Gómez, White, and Wulder 2016). However, cloud-free historical imagery from workhorse satellites like those in the Landsat program remains relatively sparse. This circumstance has forced researchers to compromise between (among other things) resolution and extent in both temporal and spatial domains (Lunetta et al. 2004). Classified maps derived from imagery at a medium spatial resolution typically possess coarse temporal resolution over a single scene (Dou and Chen 2017; Fenta et al. 2017) or over multi-scene extents (Gong et al. 2013; Sun et al. 2017), or fine temporal but coarse spatial resolution (He, Lee, and Warner 2017; Xu, Zhang, and Lin 2018). Most recently, studies have increasingly sought to create medium spatial resolution land cover time series at an annual temporal resolution, though these products may be limited in thematic resolution and spatial extent (Li, Gong, and Liang 2015; Song et al. 2016; Zhang and Weng 2016).

To mitigate the impact of limited scene availability as well as data gaps (e.g., due to failure of the scan line corrector of the ETM+ sensor), researchers have increasingly

employed data fusion for multi-temporal classifications (Gómez, White, and Wulder 2016). One popular approach to ensure spatio-temporally consistent imagery, especially for large-area classifications in heavily-clouded or undersampled regions, is the generation of best-available-pixels (BAP) composites for a given time period (White et al. 2014). Other approaches include data blending methods whereby data gaps are interpolated using temporally proximate imagery (Yin et al. 2017), as well as multi-sensor data fusion for the production of synthetic images with high temporal precision for a given date (Gong et al. 2013; Zhu et al. 2015). In this study, where annual classification accuracy was prioritized over subannual temporal precision compositing, gap-filling, and multi-date data fusion were performed at the classification stage. Using all acceptable imagery within the calendar year, classifiers were parameterized with the original reflectance retrievals and benefit from the added information content of multi-seasonal imagery, while reducing the impact of any one image on classification results. Because data fusion occurs at the classification stage (and not preceding it), pixel-wise uncertainties can be readily derived from the posterior membership probabilities of the ensemble prediction.

Despite the performance of AASG, that ensures that each automated training set was adapted to the radiometric idiosyncrasies of each new scene, and robust nonparametric classifiers like RF, numerous factors remain to affect the accuracy and consistency of land cover classifications derived from spectral data (Gray and Song 2013). Classification errors due to signal noise from subpixel heterogeneity and bidirectional reflectance distribution function effects, atmospheric contamination, as well as classifier confusion among cover classes tend to manifest in space (Song et al. 2015). At the same time, inconsistent surface reflectance retrievals due to varying specifications among sensors, sensor degradation through time, radiometric and atmospheric changes between images, as well as geolocational misalignment between dates may result in temporal inconsistencies along the classification time series (Roy et al. 2016). Spatial and temporal classification errors may then, in turn, propagate in multi-temporal classifications. To ensure greater spatio-temporal consistency in dense land cover map time series, post-classification stabilization of time series results is a critical step for improving classification accuracy and consistency (Li et al. 2014; Lu and Weng 2007). Rule-based filtering techniques based on the spatio-temporal context of a focal pixel are highly efficient for processing very large classification time series (He, Lee, and Warner 2017; Wang et al. 2015; Pouliot et al. 2014; Gao et al. 2012), while more computationally-intensive stochastic model-based approaches allow for uncertainty estimates to propagate through all steps (Wang et al. 2015; Liu and Cai 2012).

5. Conclusion

In this study, we developed an innovative automated classification algorithm that takes advantage of the synergistic value of all acceptable Landsat images in a single year, using aggregate votes from the posterior predictive distributions of multiple image composites to mitigate against misclassifications in any one image in the annual stack, and fill gaps due to missing and contaminated data, such as those from clouds and cloud shadows. Using this ensemble classification algorithm, we produced a multi-scene, annual land cover time series characterizing 21 years of dynamic land cover change

trends in the 35,000 km² Greater Houston area. Importantly, all input data were constrained to their corresponding calendar year to ensure temporal precision sufficient for researchers seeking a land cover dataset from which to investigate higher-order patterns in human–environment interactions. Land cover products of fine spatio-temporal resolution provide the means to isolate specific drivers of regional change (including environmental disturbances, economic cycles, and policy feedbacks) from their observable footprint on the ground. Furthermore, they provide sufficient temporal detail from which to estimate periodicity and temporal lags for parametrizing forecast models of future development. For this study, ecological categories were designed to be sufficiently broad to allow for temporal consistency within the hierarchical classification scheme, but still readily supplemented with the most up-to-date spatial distributions of, for instance, ecological transitions, biomass estimates, and wetland delineations that are otherwise beyond the scope of the current study.

Rapid and vast urbanization trends, coupled with more frequent and intense hurricanes, could have devastating consequences for cities like Houston in the coming decades, and especially for their most vulnerable inhabitants. Planning for these contingencies will, at the regional scale, require a concerted effort to ensure that resistance and resilience is built into future development plans. Continued advances in near-continuous, wall-to-wall Earth observation and automated land cover characterization will provide planners and policy-makers the requisite tools to make informed choices.

Acknowledgments

The authors thank the Houston Endowment, the Kinder Institute for Urban Research, and the Rice University Academy of Fellows for support of this research. DigitalGlobe data were provided by NASA's Commercial Archive Data (cad4nasa.gsfc.nasa.gov) under the National Geospatial-Intelligence Agency's NextView license agreement. We would also like to thank Eric Smith and the Kinder Institute Urban Data Platform team.

Disclosure statement

No potential conflict of interest was reported by the authors.

Data availability statement

The data that support the findings of this study are openly available at the Kinder Institute for Urban Research Urban Data Platform: www.kinderudp.org/#/datasetCatalog/zbn96g5x658z

ORCID

C.R. Hakkenberg  <http://orcid.org/0000-0002-6579-5954>

References

- Belgiu, M., and L. Drăgu. 2016. "Random Forest in Remote Sensing: A Review of Applications and Future Directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24–31. doi:[10.1016/j.isprsjprs.2016.01.011](https://doi.org/10.1016/j.isprsjprs.2016.01.011).

- Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32. doi:10.1023/A:1010933404324.
- Brody, S., H. Kim, and J. Gunn. 2013. "Examining the Impacts of Development Patterns on Flooding on the Gulf of Mexico Coast." *Urban Studies* 50 (4): 789–806. doi:10.1177/0042098012448551.
- Bullock, A., and M. Acreman. 2003. "The Role of Wetlands in the Hydrological Cycle." *Hydrology and Earth System Sciences* 7: 358–389. doi:10.5194/hess-7-358-2003.
- Cai, S., D. Liu, D. Sulla-Menashe, and M. A. Friedl. 2014. "Enhancing MODIS Land Cover Product with a Spatial-Temporal Modeling Algorithm." *Remote Sensing of Environment* 147: 243–255. Elsevier Inc. doi:10.1016/j.rse.2014.03.012.
- Chen, C., A. Liaw, and L. Breiman. 2004. "Using Random Forest to Learn Imbalanced Data." *Journal of Machine Learning Research*, No 666: 1–12.
- Congalton, R. G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37 (1): 35–46. doi:10.1016/0034-4257(91)90048-B.
- Dannenber, M. P., C. R. Hakkenberg, and C. Song. 2016. "Consistent Classification of Landsat Time Series with an Improved Automatic Adaptive Signature Generalization Algorithm." *Remote Sensing* 8: 8. doi:10.3390/rs8080691.
- Dannenber, M. P., C. Song, and C. R. Hakkenberg. 2018. "A Long-Term, Consistent Land Cover Database for the Southeastern United States Using Automatic Adaptive Signature Generalization (AASG)." *Photogrammetric Engineering & Remote Sensing* 84 (9): 35–44.
- Dou, P., and Y. Chen. 2017. "Dynamic Monitoring of Land-Use/Land-Cover Change and Urban Expansion in Shenzhen Using Landsat Imagery from 1988 to 2015." *International Journal of Remote Sensing* 38 (19): 5388–5407. doi:10.1080/01431161.2017.1339926.
- Emanuel, K. 2017. "Assessing the Present and Future Probability of Hurricane Harvey's Rainfall." *Proceedings of the National Academy of Sciences* 201716222. doi:10.1073/pnas.1716222114.
- Fenta, A. A., H. Yasuda, N. Haregeweyn, A. S. Belay, Z. Hadush, M. A. Gebremedhin, and G. Mekonnen. 2017. "The Dynamics of Urban Expansion and Land Use/Land Cover Changes Using Remote Sensing and Spatial Metrics: The Case of Mekelle City of Northern Ethiopia." *International Journal of Remote Sensing* 38 (14): 4107–4129. Taylor & Francis. doi:10.1080/01431161.2017.1317936.
- Foody, G. M. 2002. "Status of Land Cover Classification Accuracy Assessment." *Remote Sensing of Environment* 80 (1): 185–201. doi:10.1016/S0034-4257(01)00295-4.
- Fry, J. A., G. Xian, S. Jin, J. A. Dewitz, C. G. Homer, L. Yang, C. A. Barnes, N. D. Herold, and J. D. Wickham. 2011. "Completion of the 2006 National Land Cover Database for the Conterminous United States." *Photogrammetric Engineering and Remote Sensing* 77 (9): 566–588.
- Galster, G., R. Hanson, M. R. Ratcliffe, H. Wolman, S. Coleman, and J. Freihage. 2001. "Wrestling Sprawl to the Ground: Defining and Measuring an Elusive Concept." *Housing Policy Debate* 12 (4): 681–717. doi:10.1080/10511482.2001.9521426.
- Gao, F., E. B. De Colstoun, R. Ma, Q. Weng, J. G. Masek, J. Chen, Y. Pan, and C. Song. 2012. "Mapping Impervious Surface Expansion Using Medium-Resolution Satellite Image Time Series: A Case Study in the Yangtze River Delta, China." *International Journal of Remote Sensing* 33 (24): 7609–7628. doi:10.1080/01431161.2012.700424.
- Gislason, P. O., J. A. Benediktsson, and J. R. Sveinsson. 2006. "Random Forests for Land Cover Classification." *Pattern Recognition Letters* 27 (4): 294–300. doi:10.1016/j.patrec.2005.08.011.
- Gómez, C., J. C. White, and M. A. Wulder. 2016. "Optical Remotely Sensed Time Series Data for Land Cover Classification: A Review." *ISPRS Journal of Photogrammetry and Remote Sensing* 116: 55–72. doi:10.1016/j.isprsjprs.2016.03.008.
- Gong, P., J. Wang, L. Yu, Y. Zhao, Y. Zhao, L. Liang, Z. Niu, et al. 2013. "Finer Resolution Observation and Monitoring of Global Land Cover: First Mapping Results with Landsat TM and ETM+ Data." *International Journal of Remote Sensing* 34 (7): 2607–2654. doi:10.1080/01431161.2012.748992.
- Gray, J., and C. Song. 2013. "Consistent Classification of Image Time Series with Automatic Adaptive Signature Generalization." *Remote Sensing of Environment* 134 (July): 333–341. Elsevier Inc. doi:10.1016/j.rse.2013.03.022.
- Hamilton, N. 2015. "Smoother: Functions Relating to the Smoothing of Numerical Data." R Package Version 1.1. <https://cran.r-project.org/package=smoother>

- He, Y., E. Lee, and T. A. Warner. 2017. "A Time Series of Annual Land Use and Land Cover Maps of China from 1982 to 2013 Generated Using AVHRR GIMMS NDVI3g Data." *Remote Sensing of Environment* 199 (September): 201–217. Elsevier Inc. doi:10.1016/j.rse.2017.07.010.
- Herold, M., L. See, N. E. Tsensbazar, and S. Fritz. 2016. "Towards an Integrated Global Land Cover Monitoring and Mapping System." *Remote Sensing* 8 (12): 1–11. doi:10.3390/rs8121036.
- HGAC. 2018. "Houston-Galveston Area Council." Accessed 10 July 2017. <http://www.h-gac.com>
- Hijmans, R. J. 2017. "Raster: Geographic Data Analysis and Modeling." R Package Version 2.6–7. <https://cran.r-project.org/package=raster>
- Homer, C., J. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N. Herold, J. Wickham, and K. Megown. 2015. "Completion of the 2011 National Land Cover Database for the Conterminous United States-Representing a Decade of Land Cover Change Information." *Photogrammetric Engineering and Remote Sensing* 81 (5): 345–354. doi:10.14358/PERS.81.5.345.
- Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. Mckerrow, J. Nick Vandriel, and J. Wickham. 2007. "Completion of the 2001 National Land Cover Database for the Conterminous United States." *Photogrammetric Engineering & Remote Sensing* 73 (4): 337–341.
- Jaret, C., R. Ghadge, L. W. Reid, and R. M. Adelman. 2009. "The Measurement of Suburban Sprawl: An Evaluation." *City and Community*. doi:10.1111/j.1540-6040.2009.01270.x.
- Jensen, J. R., and D. C. Cowen. 1999. "Remote Sensing of Urban Suburban Infrastructure and Socio-Economic Attributes." *Photogrammetric Engineering and Remote Sensing* 65 (5): 611–622.
- Kinder Institute. 2018. "Urban Data Platform." <https://www.kinderudp.org/>.
- Knutson, T. R., J. L. McBride, J. Chan, K. Emanuel, G. Holland, C. Landsea, I. Held, J. P. Kossin, A. K. Srivastava, and M. Sugi. 2010. "Tropical Cyclones and Climate Change." *Nature Geoscience* 3: 157–163. doi:10.1038/ngeo779.
- Li, M., S. Zang, B. Zhang, S. Li, and C. Wu. 2014. "A Review of Remote Sensing Image Classification Techniques: The Role of Spatio-Contextual Information." *European Journal of Remote Sensing* 47 (1): 389–411. doi:10.5721/EuJRS20144723.
- Li, X., P. Gong, and L. Liang. 2015. "A 30-Year (1984–2013) Record of Annual Urban Dynamics of Beijing City Derived from Landsat Data." *Remote Sensing of Environment* 166: 78–90. doi:10.1016/j.rse.2015.06.007.
- Liaw, A., and M. Wiener. 2002. "Classification and Regression by RandomForest." *R News* 2 (3): 18–22.
- Liu, D., and S. Cai. 2012. "A Spatial-Temporal Modeling Approach to Reconstructing Land-Cover Change Trajectories from Multi-Temporal Satellite Imagery." *Annals of the Association of American Geographers* 102 (6): 1329–1347. doi:10.1080/00045608.2011.596357.
- Loveland, T. R., and J. L. Dwyer. 2012. "Landsat: Building a Strong Future." *Remote Sensing of Environment* 122: 22–29. doi:10.1016/j.rse.2011.09.022.
- Lu, D., and Q. Weng. 2007. "A Survey of Image Classification Methods and Techniques for Improving Classification Performance." *International Journal of Remote Sensing* 28 (5): 823–870. doi:10.1080/01431160600746456.
- Lunetta, R. S., D. M. Johnson, J. G. Lyon, and J. Crotnell. 2004. "Impacts of Imagery Temporal Frequency on Land-Cover Change Detection Monitoring." *Remote Sensing of Environment* 89 (4): 444–454. doi:10.1016/j.rse.2003.10.022.
- Maxwell, A. E., T. A. Warner, and F. Fang. 2018. "Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review." *International Journal of Remote Sensing* 39 (9): 2784–2817. Taylor & Francis. doi:10.1080/01431161.2018.1433343.
- NOAA. 2018. "U.S. Billion-Dollar Weather & Climate Disasters 1980–2017." Accessed 10 January 2018. <https://www.ncdc.noaa.gov/billions/events.pdf>
- NOAA C-CAP. 2011. *Coastal Change Analysis Program (C-CAP) Regional Land Cover*. Charleston, SC: NOAA Office for Coastal Management. Accessed March 2018. www.coast.noaa.gov/ccapftp.
- Olofsson, P., G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder. 2014. "Good Practices for Estimating Area and Assessing Accuracy of Land Change." *Remote Sensing of Environment* 148: 42–57. doi:10.1016/j.rse.2014.02.015.
- Pontius, R. G., and M. Millones. 2011. "Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment." *International Journal of Remote Sensing* 32: 4407–4429. doi:10.1080/01431161.2011.552923.

- Pouliot, D., R. Latifovic, N. Zabcic, L. Guindon, and I. Olthof. 2014. "Development and Assessment of a 250m Spatial Resolution MODIS Annual Land Cover Time Series (2000–2011) for the Forest Region of Canada Derived from Change-Based Updating." *Remote Sensing of Environment* 140: 731–743. Elsevier B.V. doi:10.1016/j.rse.2013.10.004.
- Qian, Z. 2010. "Without Zoning: Urban Development and Land Use Controls in Houston." *Cities* 27 (1): 31–41. doi:10.1016/j.cities.2009.11.006.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roy, D. P., V. Kovalsky, H. K. Zhang, E. F. Vermote, L. Yan, S. S. Kumar, and A. Egorov. 2016. "Characterization of Landsat-7 to Landsat-8 Reflective Wavelength and Normalized Difference Vegetation Index Continuity." *Remote Sensing of Environment* 185: 57–70. doi:10.1016/j.rse.2015.12.024.
- Sexton, J. O., D. L. Urban, M. J. Donohue, and C. Song. 2013. "Long-Term Land Cover Dynamics by Multi-Temporal Classification across the Landsat-5 Record." *Remote Sensing of Environment* 128: 246–258. doi:10.1016/j.rse.2012.10.010.
- Song, C., J. M. Chen, T. Hwang, A. Gonsamo, H. Croft, Q. Zhang, M. Dannenberg, Y. Zhang, C. R. Hakkenberg, and J. Li. 2015. "Ecological Characterization of Vegetation Using Multi-Sensor Remote Sensing in the Solar Reflective Spectrum." In *Remote Sensing Handbook, Vol 2. Land Resources: Monitoring, Modeling, and Mapping*, edited by P. S. Thenkabail, 533–575. London, UK: Taylor and Francis.
- Song, C., and C. E. Woodcock. 2003. "Monitoring Forest Succession with Multitemporal Landsat Images: Factors of Uncertainty." *IEEE Transactions on Geoscience and Remote Sensing* 41 (11): 2557–2567. doi:10.1109/TGRS.2003.818367.
- Song, C., C. E. Woodcock, K. C. Seto, M. P. Lenney, and S. A. Macomber. 2001. "Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects?" *Remote Sensing of Environment* 75 (2): 230–244. doi:10.1016/S0034-4257(00)00169-3.
- Song, X. P., J. O. Sexton, C. Huang, S. Channan, and J. R. Townshend. 2016. "Characterizing the Magnitude, Timing and Duration of Urban Growth from Time Series of Landsat-Based Estimates of Impervious Cover." *Remote Sensing of Environment* 175: 1–13. doi:10.1016/j.rse.2015.12.027.
- Stehman, S. V. 2013. "Estimating Area from an Accuracy Assessment Error Matrix." *Remote Sensing of Environment* 132: 202–211. doi:10.1016/j.rse.2013.01.016.
- Sun, Y., X. Zhang, Y. Zhao, and Q. Xin. 2017. "Monitoring Annual Urbanization Activities in Guangzhou Using Landsat Images (1987–2015)." *International Journal of Remote Sensing* 38 (5): 1258–1276. doi:10.1080/01431161.2016.1268283.
- U.S. Census Bureau. 2018. "Resident Population for Houston-The Woodlands-Sugar Land, TX (MSA) Retrieved from FRED, Federal Reserve Bank of St. Louis." Accessed 18 May 2018. <https://fred.stlouisfed.org/series/TX>.
- Vogelmann, J. E., A. L. Gallant, H. Shi, and Z. Zhu. 2016. "Perspectives on Monitoring Gradual Change across the Continuity of Landsat Sensors Using Time-Series Data." *Remote Sensing of Environment* 185: 258–270. Elsevier B.V. doi:10.1016/j.rse.2016.02.060.
- Wang, J., Y. Zhao, C. Li, L. Yu, D. Liu, and P. Gong. 2015. "Mapping Global Land Cover in 2001 and 2010 with Spatial-Temporal Consistency at 250m Resolution." *ISPRS Journal of Photogrammetry and Remote Sensing* 103: 38–47. doi:10.1016/j.isprsjprs.2014.03.007.
- Wang, P., C. Huang, and E. B. de Colstoun. 2017. "Mapping 2000–2010 Impervious Surface Change in India Using Global Land Survey Landsat Data." *Remote Sensing* 9 (4): 366. doi:10.3390/rs9040366.
- Weng, Q. 2012. "Remote Sensing of Impervious Surfaces in the Urban Areas: Requirements, Methods, and Trends." *Remote Sensing of Environment* 117: 34–49. Elsevier Inc. doi:10.1016/j.rse.2011.02.030.
- White, J. C., M. A. Wulder, G. W. Hobart, J. E. Luther, T. Hermosilla, P. Griffiths, N. C. Coops, et al. 2014. "Pixel-Based Image Compositing for Large-Area Dense Time Series Applications and Science." *Canadian Journal of Remote Sensing* 40 (3): 192–212. doi:10.1080/07038992.2014.945827.

- Wickham, J., S. V. Stehman, L. Gass, J. A. Dewitz, D. G. Sorenson, B. J. Granneman, R. V. Poss, and L. A. Baer. 2017. "Thematic Accuracy Assessment of the 2011 National Land Cover Database (NLCD)." *Remote Sensing of Environment* 191: 328–341. Elsevier Inc. doi:[10.1016/j.rse.2016.12.026](https://doi.org/10.1016/j.rse.2016.12.026).
- Wilkinson, G. G. 2005. "Results and Implications of a Study of Fifteen Years of Satellite Image Classification Experiments." *IEEE Transactions on Geoscience and Remote Sensing* 43 (3): 433–440. doi:[10.1109/TGRS.2004.837325](https://doi.org/10.1109/TGRS.2004.837325).
- Wilson, S. G., D. A. Plane, P. J. Mackun, T. R. Fischetti, and J. Goworowska. 2012. "Patterns of Metropolitan and Micropolitan Population Change: 2000 to 2010." *Report Number: C2010SR-01*. Washington DC.
- Woodcock, C. E., and S. Gopal. 2000. "Fuzzy Set Theory and Thematic Maps: Accuracy and Area Estimation." *International Journal of Geographical Information Science* 14 (2): 2. doi:[10.1080/136588100240895](https://doi.org/10.1080/136588100240895).
- Xu, R., H. Zhang, and H. Lin. 2018. "Annual Dynamics of Impervious Surfaces at City Level of Pearl River Delta Metropolitan." *International Journal of Remote Sensing* 39 (11): 3537–3555. doi:[10.1080/01431161.2018.1444290](https://doi.org/10.1080/01431161.2018.1444290).
- Yin, G., G. Mariethoz, Y. Sun, and M. F. McCabe. 2017. "A Comparison of Gap-Filling Approaches for Landsat-7 Satellite Data." *International Journal of Remote Sensing* 38 (23): 6653–6679. Taylor & Francis. doi:[10.1080/01431161.2017.1363432](https://doi.org/10.1080/01431161.2017.1363432).
- Yu, L., L. Liang, J. Wang, Y. Zhao, Q. Cheng, L. Hu, S. Liu, et al. 2014. "Meta-Discoveries from a Synthesis of Satellite-Based Land-Cover Mapping Research." *International Journal of Remote Sensing* 35 (13): 4573–4588. doi:[10.1080/01431161.2014.930206](https://doi.org/10.1080/01431161.2014.930206).
- Yu, W., and W. Zhou. 2017. "The Spatiotemporal Pattern of Urban Expansion in China: A Comparison Study of Three Urban Megaregions." *Remote Sensing* 9 (1): 19–21. doi:[10.3390/rs9010045](https://doi.org/10.3390/rs9010045).
- Zhang, L., and Q. Weng. 2016. "Annual Dynamics of Impervious Surface in the Pearl River Delta, China, from 1988 to 2013, Using Time Series Landsat Imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 113: 86–96. doi:[10.1016/j.isprsjprs.2016.01.003](https://doi.org/10.1016/j.isprsjprs.2016.01.003).
- Zhu, Z., S. Wang, and C. E. Woodcock. 2015. "Improvement and Expansion of the Fmask Algorithm: Cloud, Cloud Shadow, and Snow Detection for Landsats 4–7, 8, and Sentinel 2 Images." *Remote Sensing of Environment* 159: 269–277. doi:[10.1016/j.rse.2014.12.014](https://doi.org/10.1016/j.rse.2014.12.014).
- Zhu, Z., C. E. Woodcock, C. Holden, and Z. Yang. 2015. "Generating Synthetic Landsat Images Based on All Available Landsat Data: Predicting Landsat Surface Reflectance at Any Given Time." *Remote Sensing of Environment* 162: 67–83. doi:[10.1016/j.rse.2015.02.009](https://doi.org/10.1016/j.rse.2015.02.009).

Appendix 1. Landsat image list

Satellite/Sensor	Row/Path	Year	Date	Satellite/Sensor	Row/Path	Year	Date
Landsat 5 TM	25039	1997	25-Jan	Landsat 5 TM	26039	2007	23-Jul
Landsat 5 TM	25040	1997	25-Jan	Landsat 5 TM	26039	2007	8-Aug
Landsat 5 TM	26040	1997	21-Mar	Landsat 5 TM	26039	2007	24-Aug
Landsat 5 TM	25040	1997	4-Jul	Landsat 5 TM	25039	2007	4-Oct
Landsat 5 TM	25039	1997	5-Aug	Landsat 5 TM	25040	2007	4-Oct
Landsat 5 TM	26039	1997	29-Sep	Landsat 5 TM	25039	2008	12-Mar
Landsat 5 TM	26040	1997	15-Oct	Landsat 5 TM	25040	2008	12-Mar
Landsat 5 TM	26039	1997	16-Nov	Landsat 5 TM	26039	2008	19-Mar
Landsat 5 TM	26039	1997	18-Dec	Landsat 5 TM	26040	2008	19-Mar
Landsat 5 TM	26040	1997	18-Dec	Landsat 5 TM	25039	2008	29-Apr
Landsat 5 TM	25039	1997	27-Dec	Landsat 5 TM	25040	2008	29-Apr
Landsat 5 TM	25040	1997	27-Dec	Landsat 5 TM	25039	2008	4-Sep
Landsat 5 TM	26040	1998	19-Jan	Landsat 5 TM	25040	2008	4-Sep
Landsat 5 TM	26039	1998	20-Feb	Landsat 5 TM	26039	2008	27-Sep
Landsat 5 TM	26040	1998	20-Feb	Landsat 5 TM	26040	2008	27-Sep
Landsat 5 TM	25039	1998	1-Mar	Landsat 5 TM	26039	2008	29-Oct
Landsat 5 TM	25040	1998	1-Mar	Landsat 5 TM	26040	2008	30-Nov
Landsat 5 TM	25040	1998	17-Mar	Landsat 5 TM	26039	2009	2-Feb
Landsat 5 TM	26039	1998	9-Apr	Landsat 5 TM	25039	2009	11-Feb
Landsat 5 TM	26040	1998	9-Apr	Landsat 5 TM	26039	2009	7-Apr
Landsat 5 TM	25039	1998	7-Jul	Landsat 5 TM	26040	2009	7-Apr
Landsat 5 TM	26039	1998	3-Nov	Landsat 5 TM	25039	2009	18-May
Landsat 5 TM	25040	1998	14-Dec	Landsat 5 TM	25040	2009	18-May
Landsat 5 TM	25039	1998	30-Dec	Landsat 5 TM	25039	2009	22-Aug
Landsat 5 TM	25040	1999	31-Jan	Landsat 5 TM	25040	2009	22-Aug
Landsat 5 TM	26039	1999	28-Apr	Landsat 5 TM	26040	2009	1-Nov
Landsat 5 TM	25039	1999	7-May	Landsat 5 TM	25040	2009	10-Nov
Landsat 5 TM	25040	1999	7-May	Landsat 5 TM	26039	2009	3-Dec
Landsat 5 TM	25039	1999	27-Aug	Landsat 5 TM	26040	2009	3-Dec
Landsat 5 TM	26039	1999	19-Sep	Landsat 5 TM	25039	2010	18-Mar
Landsat 5 TM	26040	1999	19-Sep	Landsat 5 TM	25040	2010	18-Mar
Landsat 5 TM	26040	1999	21-Oct	Landsat 5 TM	26039	2010	25-Mar
Landsat 5 TM	25040	1999	15-Nov	Landsat 5 TM	26040	2010	25-Mar
Landsat 7 ETM+	26040	1999	30-Nov	Landsat 5 TM	25039	2010	5-May
Landsat 5 TM	25039	1999	1-Dec	Landsat 5 TM	25040	2010	5-May
Landsat 5 TM	26039	1999	24-Dec	Landsat 5 TM	26039	2010	31-Jul
Landsat 7 ETM+	26039	2000	17-Jan	Landsat 5 TM	25039	2010	25-Aug

Landsat 5 TM	25039	2000	3-Feb	Landsat 5 TM	25040	2010	25-Aug
Landsat 5 TM	25040	2000	3-Feb	Landsat 5 TM	26040	2010	3-Oct
Landsat 7 ETM+	26039	2000	12-Aug	Landsat 5 TM	26039	2010	4-Nov
Landsat 7 ETM+	26040	2000	12-Aug	Landsat 5 TM	26040	2010	6-Dec
Landsat 5 TM	25040	2000	13-Aug	Landsat 5 TM	26039	2011	8-Feb
Landsat 7 ETM+	26039	2000	29-Sep	Landsat 5 TM	26040	2011	8-Feb
Landsat 7 ETM+	26040	2000	29-Sep	Landsat 5 TM	26040	2011	29-Apr
Landsat 5 TM	25039	2000	30-Sep	Landsat 5 TM	26039	2011	15-May
Landsat 5 TM	25040	2000	30-Sep	Landsat 5 TM	26039	2011	3-Aug
Landsat 7 ETM+	26040	2000	16-Dec	Landsat 5 TM	26040	2011	3-Aug
Landsat 5 TM	25039	2000	19-Dec	Landsat 5 TM	25039	2011	28-Aug
Landsat 5 TM	25039	2001	4-Jan	Landsat 5 TM	25040	2011	28-Aug
Landsat 5 TM	25040	2001	20-Jan	Landsat 5 TM	25039	2011	13-Sep
Landsat 7 ETM+	26039	2001	4-Feb	Landsat 5 TM	25040	2011	13-Sep
Landsat 7 ETM+	26040	2001	4-Feb	Landsat 5 TM	25039	2011	31-Oct
Landsat 5 TM	25040	2001	5-Feb	Landsat 5 TM	25040	2011	31-Oct
Landsat 5 TM	25039	2001	26-Apr	Landsat 7 ETM+	26040	2012	2-Jan
Landsat 5 TM	26039	2001	22-Jul	Landsat 7 ETM+	25039	2012	11-Jan
Landsat 5 TM	26040	2001	22-Jul	Landsat 7 ETM+	25040	2012	11-Jan
Landsat 5 TM	26039	2001	26-Oct	Landsat 7 ETM+	26039	2012	18-Jan
Landsat 5 TM	26040	2001	26-Oct	Landsat 7 ETM+	25039	2012	27-Jan
Landsat 5 TM	25039	2001	4-Nov	Landsat 7 ETM+	25040	2012	27-Jan
Landsat 5 TM	25040	2001	4-Nov	Landsat 7 ETM+	26039	2012	22-Mar
Landsat 5 TM	26039	2002	14-Jan	Landsat 7 ETM+	26040	2012	22-Mar
Landsat 5 TM	26040	2002	14-Jan	Landsat 7 ETM+	26040	2012	7-Apr
Landsat 5 TM	25039	2002	24-Feb	Landsat 7 ETM+	26039	2012	23-Apr
Landsat 5 TM	25040	2002	24-Feb	Landsat 7 ETM+	25039	2012	18-May
Landsat 5 TM	25039	2002	3-Aug	Landsat 7 ETM+	25040	2012	18-May
Landsat 5 TM	25040	2002	3-Aug	Landsat 7 ETM+	26039	2012	29-Aug
Landsat 5 TM	26039	2002	27-Sep	Landsat 7 ETM+	26040	2012	29-Aug
Landsat 5 TM	26040	2002	27-Sep	Landsat 7 ETM+	25039	2012	23-Sep
Landsat 5 TM	26039	2002	29-Oct	Landsat 7 ETM+	25040	2012	23-Sep
Landsat 5 TM	26040	2002	29-Oct	Landsat 7 ETM+	26039	2012	17-Nov
Landsat 5 TM	25039	2002	7-Nov	Landsat 7 ETM+	26040	2012	17-Nov
Landsat 5 TM	25040	2002	23-Nov	Landsat 7 ETM+	25039	2012	12-Dec
Landsat 5 TM	26040	2003	1-Jan	Landsat 7 ETM+	25040	2012	12-Dec
Landsat 5 TM	26039	2003	17-Jan	Landsat 8 OLI	26039	2013	27-Mar
Landsat 5 TM	25039	2003	31-Mar	Landsat 8 OLI	26040	2013	27-Mar
Landsat 5 TM	25040	2003	31-Mar	Landsat 8 OLI	25040	2013	11-Apr
Landsat 5 TM	25039	2003	18-May	Landsat 8 OLI	26040	2013	4-May

Landsat 5 TM	25040	2003	18-May	Landsat 8 OLI	25039	2013	13-May
Landsat 5 TM	25040	2003	7-Sep	Landsat 8 OLI	26039	2013	8-Aug
Landsat 5 TM	26039	2003	30-Sep	Landsat 8 OLI	25039	2013	17-Aug
Landsat 5 TM	26040	2003	30-Sep	Landsat 8 OLI	25040	2013	17-Aug
Landsat 5 TM	25039	2003	10-Nov	Landsat 8 OLI	26039	2013	25-Sep
Landsat 5 TM	26039	2003	19-Dec	Landsat 8 OLI	26040	2013	25-Sep
Landsat 5 TM	26040	2003	19-Dec	Landsat 8 OLI	25039	2013	23-Dec
Landsat 5 TM	26039	2004	21-Feb	Landsat 8 OLI	25040	2013	23-Dec
Landsat 5 TM	26040	2004	8-Mar	Landsat 8 OLI	26039	2014	15-Jan
Landsat 5 TM	26040	2004	9-Apr	Landsat 8 OLI	26040	2014	15-Jan
Landsat 5 TM	25039	2004	4-May	Landsat 8 OLI	25039	2014	13-Mar
Landsat 5 TM	25040	2004	4-May	Landsat 8 OLI	25040	2014	13-Mar
Landsat 5 TM	26039	2004	14-Jul	Landsat 8 OLI	26039	2014	20-Mar
Landsat 5 TM	26040	2004	14-Jul	Landsat 8 OLI	26040	2014	20-Mar
Landsat 5 TM	26039	2004	15-Aug	Landsat 8 OLI	25039	2014	16-May
Landsat 5 TM	25039	2004	9-Sep	Landsat 8 OLI	25040	2014	16-May
Landsat 5 TM	25040	2004	9-Sep	Landsat 8 OLI	26039	2014	14-Oct
Landsat 5 TM	25039	2004	14-Dec	Landsat 8 OLI	26040	2014	14-Oct
Landsat 5 TM	25040	2004	14-Dec	Landsat 8 OLI	25039	2014	24-Nov
Landsat 5 TM	26039	2005	11-Mar	Landsat 8 OLI	25040	2014	24-Nov
Landsat 7 ETM+	25040	2005	12-Mar	Landsat 8 OLI	25039	2015	27-Jan
Landsat 5 TM	26040	2005	12-Apr	Landsat 8 OLI	25040	2015	27-Jan
Landsat 5 TM	25039	2005	23-May	Landsat 8 OLI	25040	2015	12-Feb
Landsat 5 TM	25040	2005	24-Jun	Landsat 8 OLI	26039	2015	23-Mar
Landsat 5 TM	25039	2005	10-Jul	Landsat 8 OLI	26040	2015	23-Mar
Landsat 5 TM	25040	2005	26-Jul	Landsat 8 OLI	26039	2015	1-Oct
Landsat 5 TM	26039	2005	21-Oct	Landsat 8 OLI	26040	2015	1-Oct
Landsat 5 TM	26040	2005	21-Oct	Landsat 8 OLI	25039	2015	10-Oct
Landsat 5 TM	26039	2005	22-Nov	Landsat 8 OLI	25040	2015	10-Oct
Landsat 5 TM	26040	2005	22-Nov	Landsat 8 OLI	26039	2015	4-Dec
Landsat 7 ETM+	25040	2005	23-Nov	Landsat 8 OLI	26040	2015	4-Dec
Landsat 5 TM	25039	2005	31-Dec	Landsat 8 OLI	26039	2016	6-Feb
Landsat 5 TM	25040	2005	31-Dec	Landsat 8 OLI	26040	2016	6-Feb
Landsat 5 TM	25039	2006	18-Jan	Landsat 8 OLI	25039	2016	2-Mar
Landsat 5 TM	25040	2006	18-Jan	Landsat 8 OLI	26039	2016	25-Mar
Landsat 5 TM	26039	2006	26-Feb	Landsat 8 OLI	25040	2016	3-Apr
Landsat 5 TM	26040	2006	14-Mar	Landsat 8 OLI	25039	2016	5-May
Landsat 5 TM	25040	2006	8-Apr	Landsat 8 OLI	25040	2016	5-May
Landsat 5 TM	26039	2006	17-May	Landsat 8 OLI	25040	2016	28-Oct
Landsat 5 TM	26040	2006	17-May	Landsat 8 OLI	26039	2016	20-Nov

Landsat 5 TM	25039	2006	27-Jun	Landsat 8 OLI	26040	2016	20-Nov
Landsat 5 TM	25039	2006	4-Dec	Landsat 8 OLI	26040	2016	6-Dec
Landsat 5 TM	25040	2006	4-Dec	Landsat 8 OLI	25039	2016	15-Dec
Landsat 5 TM	26039	2006	27-Dec	Landsat 8 OLI	26039	2017	7-Jan
Landsat 5 TM	26040	2006	27-Dec	Landsat 8 OLI	26040	2017	7-Jan
Landsat 5 TM	26039	2007	28-Jan	Landsat 8 OLI	26039	2017	23-Jan
Landsat 5 TM	26040	2007	28-Jan	Landsat 8 OLI	26040	2017	23-Jan
Landsat 5 TM	26040	2007	13-Feb	Landsat 8 OLI	25039	2017	21-Mar
Landsat 5 TM	26039	2007	1-Mar	Landsat 8 OLI	25040	2017	21-Mar
Landsat 5 TM	25039	2007	11-Apr	Landsat 8 OLI	26040	2017	28-Mar
Landsat 5 TM	25040	2007	11-Apr	Landsat 8 OLI	25039	2017	6-Apr
Landsat 5 TM	26040	2007	18-Apr	Landsat 8 OLI	25040	2017	6-Apr
Landsat 5 TM	25039	2007	13-May	Landsat 8 OLI	26039	2017	15-May
Landsat 5 TM	25040	2007	13-May				

Appendix 2. Validation imagery list

Sensor	Date
AndrewLonnieSikes aerial	17-Dec-98
AndrewLonnieSikes aerial	17-Dec-98
AndrewLonnieSikes aerial	17-Dec-98
AndrewLonnieSikes aerial	18-Jan-99
AndrewLonnieSikes aerial	18-Jan-99
AndrewLonnieSikes aerial	18-Jan-99
IKONOS	2-Jan-03
IKONOS	11-Nov-03
IKONOS	17-Aug-04
Quickbird	24-Apr-05
Quickbird	25-May-05
HGAC aerial	5-Aug-06
Quickbird	7-Jan-07
Quickbird	7-Jan-07
Quickbird	29-Jul-09
Quickbird	17-Jan-10
Worldview	23-Jul-10
Quickbird	24-Jul-10
Worldview	27-Jan-12
Worldview	27-Jan-12
Quickbird	30-Jul-12
Worldview	6-Jan-13
Quickbird	3-Jul-13
Worldview	15-Jan-14
Worldview	8-Jan-15
Worldview	16-Jan-15
Worldview	19-Jan-15
Worldview	28-Jan-16
Worldview	15-Jul-16
Worldview	22-Jan-17

Appendix 3. Land cover class definitions

Class code	Class name	NLCD code	NLCD class description*
1	Barren / Sand	31	Anthropogenically impacted areas of unconsolidated dirt and rock (e.g. dirt lots, strip mines, gravel pits and other accumulations of earthen material)
2	Developed - Open	21	Developed areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.
4	Developed – Low intensity	22	Developed areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20-49 percent of total cover. These areas most commonly include single-family housing units.
8	Developed – Medium intensity	23	Developed areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50-79 percent of the total cover. These areas most commonly include single-family housing units
13	Developed – High intensity	24	Highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80 to 100 percent of the total cover.
21	Cultivated Crops	82	Cultivated Crops - Areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20 percent of total vegetation. This class also includes all land being actively tilled.
35	Grassland / Pasture	71, 81, 95	Upland and wetland areas dominated by graminoid or herbaceous vegetation which may be subject to intensive management such as tilling, grazing, or the production of seed or hay crops.
50	Forest	41, 42, 43, 52, 90	Upland and wetland areas where forest (>5m) and shrubland (<5m) vegetation accounts for greater than 20 percent of vegetative cover.
71	Water	11	All areas of open water

* Derived NLCD class definitions from Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., Megown, K., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. Photogramm. Eng. Remote Sensing 81, 345–354.

Appendix 4. Developed classes agreement with NLCD maps for 2001, 2006, and 2011.

Table 1. Combined Developed classes only, where NLCD is the reference map. U_{AG} - user's agreement; P_{AG} - producer's agreement.

	U_{AG}	P_{AG}	F-score
2001	86.6%	92.4%	89.4%
2006	81.6%	97.2%	88.7%
2011	81.3%	97.2%	88.5%

Appendix 5. Distribution of validation sample set in space and time.

Table 1. Class cover as total proportion of area (%). Values represent the areal proportions for each class in all validation imagery (sample) and the entire study area (population).

Class	Validation imagery	Full HGAC
Sand / Barren	0.3%	0.3%
Dev-Open	11.0%	8.2%
Dev-Low	6.1%	2.8%
Dev-Med	6.2%	2.3%
Dev-High	5.3%	1.8%
Cultivated Crops	7.2%	12.0%
Grass	32.3%	35.5%
Forest	27.0%	26.5%
Water	4.5%	10.7%

Table 2. Total samples (pixels) by Year.

Year	Total samples
1999	259
2003	213
2004	117
2005	306
2006	100
2007	228
2009	104
2010	280
2012	298
2013	240
2014	154
2015	387
2016	272
2017	132

Appendix 6. Raw confusion matrices (samples)

Table 1. Raw sample confusion matrix (full). UA – user’s accuracy; UE – User’s error (commission); PA - producer’s accuracy; PE - producer’s error (omission); OA – overall accuracy; OE – overall error.

	Reference										
	Sand/Barren	Dev-Open	Dev-Low	Dev-Med	Dev-High	Cultivated Crops	Grassland / Pasture	Forest	Water	UA	UE
Sand / Barren	73	0	1	0	2	0	0	0	0	96.1%	4.4%
Dev-Open	1	125	102	25	10	2	19	10	2	42.2%	5.6%
Dev-Low	0	25	165	85	20	0	6	0	0	54.8%	5.6%
Dev-Med	1	13	71	208	71	0	5	0	1	56.2%	5.1%
Dev-High	2	4	18	52	247	0	0	0	0	76.5%	4.6%
Map											
Cultivated Crops	0	2	1	0	0	110	41	4	1	69.2%	7.2%
Grass	2	37	10	2	1	28	646	24	4	85.7%	2.5%
Forest	0	26	8	1	0	1	38	462	5	85.4%	3.0%
Water	3	1	2	0	1	0	10	3	196	90.7%	3.9%
PA	46%	50%	36%	40%	57%	86%	83%	92%	95%	OA	77.9%
PE	11%	33%	24%	22%	15%	13%	4%	3%	2%	OE	1.6%